SEX-SPECIFIC REGULATION OF PIRNA BIOGENESIS DURING GERMLINE DEVELOPMENT

by

Charlotte P. Choi

A dissertation submitted to The Johns Hopkins University

in conformity with the requirements for the degree of

Doctor of Philosophy

Baltimore, Maryland

March 2021

Abstract

Piwi-interacting RNAs (piRNAs) are a class of small RNAs that play an essential and conserved role in germline development and act as the frontlines of defense against foreign genetic elements. Across the animal kingdom, piRNAs share common features such as sex-specific germline enrichment and genomic clustering. The piRNA pathway have been described in the germline of both sexes for many metazoan organisms, including humans. Despite extensive characterization of the piRNA pathway, the underlying mechanism underlying sex-specific piRNA biogenesis and targeting remain elusive.

In *C. elegans*, the transcription factor SNPC-4 is essential for the biogenesis of both male and female piRNAs and localizes to subnuclear foci in both male and female meiotic germ cells, likely corresponding to the locations of the two piRNA genomic clusters. We demonstrate SNPC-4 binds sex-specific factors to regulate piRNA transcription in a sex-specific manner. Through SNPC-4 purification and mass spectrometry in feminized and masculinized *C. elegans* germlines, I have identified a novel male piRNA transcription factor, SNPC-1.3. SNPC-1.3 promotes the expression of male piRNAs during spermatogenesis but is dispensable for female piRNA expression. Moreover, the promoter of *snpc-1.3* contains three highly conserved binding sites of TRA-1, a well-documented master regulator of sex-determination. During oogenesis, TRA-1 is upregulated during oogenesis and

binds the snpc-1.3 promoter to repress its transcription, leading to the expression of female over male piRNAs.

Primary Reader: John Kim

Secondary Reader: Yumi Kim

“He understood for the first time that the world is not dumb at all, but merely waiting for someone to speak to it in a language it understands.”

-Susanna Clarke, Jonathan Strange and Mr. Norrell

Acknowledgements

I am thankful to the colleagues and mentors that have shaped my time as a graduate student. I am first grateful to my advisor, John Kim, for giving me the resources and opportunities to fully engage in the scientific process. I would also like to thank my thesis committee members, Yumi Kim, Mark Van Doren, and Carl Wu, for their wealth of technical help and scientific insight. I am grateful to Rebecca Tay, who contributed much to the scientific ideas and experimental work that were invaluable for the SNPC-1.3 project. I would like to thank Himani Galagali who has been my main confidant for scientific correspondence and whose unbounding passion for science made my time in graduate school an exciting and joyous experience. I am thankful to Natasha Weiser who was my role model throughout graduate school and helped me gather the courage to trust my instincts. I am grateful to all the undergraduate researchers who have worked with me, especially Emily Xu and Darius Mostaghimi, whose hard work and insightful questions contributed to pursuing many experimental avenues that ended up critical for several of my projects. I am also thankful to Mindy Clark, Amelia Alessi, Jessie Kirshner, Greg Fuller, Nathan Roach, and Mallory Freeberg for their scientific discussion and for making sure I did not go a day in the Kim lab without laughing. Lastly, I would like to thank Enis Afgan and the Galaxy team, who generously provided computational resources and bioinformatic expertise that help advance many of my projects.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction to sex-specific piRNAs function in the germline

Piwi-interacting RNAs (piRNAs) are a distinct class of small noncoding RNAs that function to protect germline integrity. piRNAs were first characterized in the *Drosophila* testis as a class of small RNAs that silence *Stellate* (*Ste*), a gene that resides on the X chromosome and, when inappropriately expressed in the testes, can lead to arrest during spermatogenesis (Aravin et al., 2001). Since their initial discovery, many studies have identified piRNAs in the gametes of numerous metazoan species and loss of the piRNA pathway is frequently linked to severe germline defects and sterility.

Deep sequencing of small RNAs in the germline has revealed piRNAs are remarkably diverse in sequence. In *Caenorhabditis elegans*, over 15,000 unique piRNA species are encoded in the genome (Gu et al., 2012; Ruby et al., 2006). Furthermore, the sequences of piRNAs are evolving at a rapid pace, such that piRNAs among even closely related species show no evidence of conservation (Wit et al., 2009). As a result, the piRNA pathway is strikingly versatile in function. In flies and mammals, piRNAs target and silence transposable elements and retroviruses (Aravin et al., 2007; Brennecke et al., 2007; Kuramochi-Miyagawa et al., 2008; Malone et al., 2009). During *C. elegans* germline development, piRNAs predominantly bind protein coding genes (Lee et al., 2012; Shen et al., 2018).

Finally, many piRNAs display sex-specific expression and function. Studies in *Bombyx mori* have revealed inheritance of a single piRNA can determine the sexual fate of the germline (Kiuchi et al., 2014). In male mice, piRNAs play a critical function during the epigenetic reprogramming of fetal germline stem cells, while in females, they are dispensable for fertility (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008; Malki et al., 2014; Reuter et al., 2011; Tharp et al., 2020). This introduction aims to summarize what is currently known about the central mechanisms underlying piRNA function during germline development, as well as document sex-specific differences in piRNA biogenesis and function, with particular consideration towards insect, mammalian, and nematode model systems, where piRNAs have been best characterized.

1.2 The function of piRNAs in *Drosophila melanogaster*

In *Drosophila*, mature piRNAs are 23–31 nucleotides (nt) in length and can be loaded into one of three PIWI proteins, Piwi, Aubergine (Aub), and Ago3, which are essential for piRNA stability and function. The most well-characterized function of PIWI proteins in the *Drosophila* germline is the silencing of transposons and retroviruses (Brennecke et al., 2007; Li et al., 2009; Malone et al., 2009; Vagin et al., 2006). Left unchecked, deleterious genetic elements such as transposons can wreak havoc on the genome inherited by the offspring. PIWI proteins have additional functions during spermatogenesis such as regulating endogenous gene expression, that when overexpressed leads to spermatogenic defects (Aravin et al., 2004; Nishida et al., 2007). Correspondingly, loss of any one of the three Piwi proteins in either sex leads to severe or complete loss in fertility.

1.2.1 *Drosophila* oogenesis

Developing fly oocytes, owing to their large size, are a tractable system to study piRNA function and subcellular localization during gametogenesis. As a result, piRNA biogenesis has been extensively characterized during *Drosophila* oogenesis. The piRNA pathway is active throughout all stages of oogenesis, from mitotic proliferation to meiosis, as well as in the surrounding somatic follicle cells that act to support oocyte development (Brennecke et al., 2007). In ovarian germ cells, piRNAs are initially derived from long piRNA precursors (pre-piRNAs) that are convergently transcribed by RNA Polymerase II from genomic regions called dual-strand clusters. Pre-piRNAs are then exported out to the cytoplasm and undergo endonucleolytic cleavage by Zucchini (Zuc) on the mitochondrial surface (Nishimasu et al., 2012; Ipsaro et al., 2012). Zucchini derived piRNAs are either loaded into Piwi or Aub (Han et al., 2015; Mohn et al., 2015; Senti et al., 2015; Watanabe et al., 2010), and the preference for PIWI protein determines whether a piRNA participates within one of two distinct transposon silencing mechanisms.

The loading into Piwi licenses translocation into the nucleus where Piwi binds nascent transposon transcripts and induces co-transcriptional gene silencing (Brower-Toland et al., 2007; Moshkovich and Lei, 2010). Alternatively, piRNAs can be loaded into Aub to initiate a robust post-transcriptional gene silencing program, called the ping-pong amplification cycle, in electron-dense, perinuclear nuage (Aravin et al., 2004; Brennecke et al., 2007; Lim and Kai, 2007; Wang et al., 2015). Aub-bound piRNAs, generated by Zuc, first bind and cleave transposon transcripts, resulting in two new fragments (Han et al., 2015; Ipsaro et al., 2012;

Mohn et al., 2015; Nishimasu et al., 2012). The 3' fragment is loaded into Ago3 and serves as a guide to induce another round of cleavage of piRNA precursor transcripts that are derived from dual-strand clusters. Production of Ago3 piRNAs requires Aub and is specifically enriched for an adenine 10 nt downstream from the 5' end which is a signature of Aub catalysis. In vitro work has shown Aub has a specific affinity for adenine at the position complementary to the 5' most nucleotide of Ago3 bound piRNAs (Wang et al., 2014). Aub cleaves transposon by first binding target transcripts via complementary base pairing 10 nt downstream from it's cut site. The new 5' end of complementary piRNAs are loaded into Ago3, as such Ago3 bound piRNAs are biased for an adenine at the 10 nt position due to base pairing with the complementary 5' uridine of Aub bound piRNAs. These Ago3-dependent piRNAs then direct another round of cleavage to produce another set of Aub-bound piRNAs. This reciprocal cleavage, mediated by Aub and Ago3 (Brennecke et al., 2007; Gunawardane et al., 2007), achieves amplification of piRNA expression while simultaneously inducing rapid destruction of transposons.

The piRNA pathway is active in both ovarian germ cells and somatic follicles, however, several studies have reported key differences between the two cell types. First, the piRNA pathway components used to carry out downstream functions are different between somatic follicle and ovarian germ cells. While Piwi functions in the nucleus of both cell types to trigger co-transcriptional silencing, Aub and Ago3-mediated ping-pong amplification is absent in the soma (Robine et al., 2009). The targets of somatic piRNAs are also different. Instead of targeting transposons, somatic piRNAs silence retroviruses, which are made in somatic

follicle cells and can go on to infect nearby germ cells (Brennecke et al., 2007; Li et al., 2009; Malone et al., 2009). Lastly, genomic origins of somatic and germ cell piRNAs are not the same. piRNAs enriched in ovarian germ cells predominantly originate from dual-strand clusters that are embedded in large heterochromatic regions (Klattenhoff et al., 2009). These clusters house a panoply of inactive transposons fragments that act as guides for the piRNA pathway to target active transposons encoded in other regions of the genome. Recent studies have revealed that heterochromatic marks are critical for licensing transcription at dual-strand piRNA clusters (which will be covered in further detail later on in this review) (Klattenhoff et al., 2009; Mohn et al., 2014). Somatic piRNAs are also derived from heterochromatic regions, but are transcribed from large uni-stranded clusters that do not rely solely on the local chromatin environment for transcriptional activation. Transcription of Flamenco (Flam), the largest uni-stranded piRNA cluster, requires the transcription factor Cubitus interruptus (Ci) (Goriaux et al., 2014). The Ci binding site, found within Flam, is conserved only among other Drosophila species, suggesting somatic piRNAs are rapidly evolving. Indeed, the function of somatic follicle piRNAs seems unique to fly oogenesis, as piRNAs that target retroviruses have not been observed in the fly testis or other model systems.

1.2.2 *Drosophila* spermatogenesis

Although the piRNA pathway is critical for male fertility, piRNA activity in the *Drosophila* testis is largely restricted to a small number of cell types undergoing mitosis (Fuller 1993). Spermatogenesis begins at the apical tip, where each germ line stem cell (GSCs) divides asymmetrically to produce a single GSC and a

daughter gonialblast (GB). GBs undergo 4 consecutive rounds of mitosis as well as incomplete cytokinesis to produce 16 interconnected germ cells called the spermatogonia. Spermatogonial cells (SGs) differentiate into spermatocytes (SCs), which coincides with the start of meiosis. SCs progress through two meiotic divisions while at the same time undergoing dramatic morphological changes in order to produce mature haploid spermatids that are viable for fertilization. Piwi has been found to be only expressed in the nucleus of GSCs, GBs, and in several neighboring hub and somatic cyst stem cells (CySCs) located at the apical tip of the testis (Cox et al. 2000). The hub cells maintain GSC identity at the apical tip, while the main function of CySCs is to support germline differentiation and maturation.

Piwi possesses both cell autonomous and non-autonomous functions in the adult *Drosophila* testis to promote early germ cell differentiation (Cox et al., 2000; Gonzalez et al., 2015). Specific loss of Piwi in the CySCs results in CySCs differentiation defects, increase in the number of GSCs, and a failure in GSC differentiation (Gonzalez et al., 2015). Small RNA sequencing of piRNAs bound to Piwi in testes reveal piRNAs that uniquely target Fasciclin 3, a gene required for somatic differentiation in the gonad. In addition, half of Piwi-bound piRNAs in the testes target protein coding genes (Gonzalez et al., 2015), whereas the majority of Piwi-bound piRNAs in ovaries target retroviral sequences (Brennecke et al., 2007). Piwi also plays a critical role in GSCs, as loss of Piwi in the germline also results in germ cell developmental arrest (Gonzalez et al., 2015). The specific function of Piwi bound piRNAs in the testis germ cells has not been fully characterized.

However, they likely have similar functions to piRNAs in ovarian germ cells, as a third of testicular Piwi-bound piRNAs target transposons.

Aub and Ago3 exhibit distinct expression profiles during spermatogenesis. In the female germline, Ago3 and Aub are broadly expressed throughout oogenesis and show colocalization at the perinuclear nuage. In flies, transposons are active throughout oocyte development. Therefore, Ago3 and Aub must function together continuously during oocyte development, in order to sustain ping-pong amplification and maintain a robust pool of piRNAs to sufficiently silence transposons. In the testis germline, Ago3 expression is confined to early stages of development and localizes to the perinuclear nuage of GSCs and GBs, and early stage SGs, whereas Aub shows widespread expression from GSCs to primary SCs (Nagao et al., 2010; Quénerch'du et al., 2016).

Recent studies of Ago3 and Aub function in early versus late spermatogenesis indicate that ping-pong amplification may occur predominantly during early spermatogenesis but not in late-stage spermatocytes. Correspondingly, small RNA sequencing of testes arrested at early versus late stages of spermatogenesis reveal piRNAs mapping to transposons and dual-strand clusters are more enriched during early stages of mitotic germline development than at later stages (Quénerch'du et al., 2016). In fly ovaries, the Rhino-Deadlock-Cutoff (RDC) nuclear complex is critical for the transcription of dual-strand piRNAs. Likewise, preliminary findings report that the RDC complex binds dual-strand clusters in the GSCs and SGs of the testis (Chen et al., 2020). In SCs, however, RDC complex binding at dual-strand clusters is lost. Intriguingly,

male flies that lose RDC complex function are moderately fertile, while rhi, del, or cutoff mutant females are sterile. As dual-strand cluster piRNAs act as the primary guides for ping pong cycle-dependent transposon silencing, these data suggest transposon dysregulation is not a predominant threat to testis germline integrity, especially during later meiotic stages. Additionally, a majority of Ago3- and Aub-bound targets do not map to transposons, but instead correspond to two male-specific endogenous loci, Suppressor of *Stellate Su(Ste)* and *At-chX* (Quénerch'du et al., 2016).

The *Su(Ste)* piRNA locus is critical for silencing a series of repeats of *Stellate*, a gene that encodes for the Stellate protein, which, when de-repressed, leads to severe meiotic defects (Schmidt et al., 1999). The *Su(Ste)* locus is Y chromosome-linked, and therefore male specific, such that male *Su(Ste)* mutant flies are sterile and develop large Stellate crystal aggregates in maturing SCs that are not found in wild-type testis. Unlike piRNA-dependent transposon regulation, *Su(Ste)* biogenesis and Ste repression are most likely spatiotemporally distinct, as *Stellate* transcripts are only expressed in SCs while *Su(Ste)* piRNA biogenesis factors (except Aub) are restricted to early mitotic germ cells (Aravin et al., 2004; Klattenhoff et al., 2007; Kotelnikov et al., 2009). piRNAs derived from *Su(Ste)* are transcribed bidirectionally and generate piRNA precursor transcripts that are both antisense and sense to *Su(Ste)*. The processing of mature piRNAs from nascent *Su(Ste)* transcripts require Ago3 and Aub, as well as dual-strand piRNA-precursor processing factors Armitage (Armi) and Zuc (Aravin et al., 2004; Pane et al., 2007). However, the mechanism that underlies how these piRNA factors interact to drive

Su(Ste) piRNA biogenesis is still unclear. In addition to the *Su(Ste)* locus, the second largest class of testis-derived piRNAs are from the AT-ChX locus on the X chromosome. Many At-ChX-derived piRNAs show strong complementarity to vasa mRNA, a germline-enriched transcript that has been shown to be critical for oocyte differentiation, and loss in Aub leads to increased Vasa protein expression in the fly testis (Nagao et al., 2010; Nishida et al., 2007). However, whether the At-ChX locus directly regulates Vasa protein expression during male germline development is still not known.

In summary, piRNAs play an essential role in transposon silencing, as well as the regulation of male-specific protein coding genes during fly testis development. Nevertheless, the exact mechanism that underlies piRNA biogenesis or function in the fly testis is still not well understood. Loss of Ago3 in the testis leads to a moderate reduction in ping-pong activity at transposon transcripts (Quénerch'du et al., 2016). Loss of Aub, in contrast, results in the collapse of ping-pong amplified piRNAs at transposons in a similar manner to that observed in aub mutant oocytes. In addition, ago3 mutant males are semi-fertile, while aub mutants are sterile (Li et al., 2009; Schmidt et al., 1999), suggesting Aub may be able to function alone in the testis to achieve ping-pong cycle amplification or participate in an alternative biogenesis mechanism altogether. In fly oocytes, Aub and Ago3 forms a heterodimer that is mediated by a Tudor domain containing E3 ligase, Qin. The loss of qin leads to a decoupling of Ago3 and Aub, along with promoting the formation of homotypic Aub dimers (Zhang et al., 2011). In mice, ping-pong cycle amplification is accomplished by a single Piwi protein. However,

in flies, homotypic Aub dimers are insufficient to sustain ping-pong activity in the ovarian germline (Zhang et al., 2011). These data suggest that Ago3 and Aub target transposons and other genes in the fly testis but may silence gene targets through a possibly alternative mechanism.

1.3 The Function of piRNAs in mammalian systems

piRNAs in mammalian systems are dynamically expressed throughout germline development and function in critical processes such as epigenetic reprogramming of the fetal germline, meiosis, and spermiogenesis (Aravin et al., 2008; Deng and Lin, 2002; Gou et al., 2014; Hsieh et al., 2020; Kuramochi-Miyagawa et al., 2008; Reuter et al., 2011; Vourekas et al., 2012). Mammalian piRNAs have been best characterized in mouse testis development, where they play distinct, critical roles during early mitotic and late meiotic stages. Intriguingly, the piRNA pathway is dispensable in the mouse female germline, as female mice without an active piRNA pathway are fertile and display no significant defects during oogenesis (Carmell et al., 2007; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004). However, studies of the piRNA pathway within both fetal and adult ovaries in other mammalian systems show piRNAs are active during all stages of oogenesis, suggesting female mice may be a unique exception where piRNAs do not play an essential role during germline development.

1.3.1 Overview of piRNAs during murine testes development

The mouse genome encodes three Piwi proteins, MILI, MIWI, and MIWI2, which interact with piRNAs that are 26–29 nt in length. Piwi biogenesis begins early during fetal germline development in which MILI is expressed in gonocytes. MILI

is continuously expressed throughout male gametogenesis until the round spermatid stage. MIWI2 is only expressed in gonocytes for a short period around the time of birth. In contrast, MIWI and MILI are expressed in the adult testis, from late pachytene to the spermatid stage (Kuramochi-Miyagawa et al., 2004). Piwi protein interaction is critical for piRNA stability, thus the characterization of each Piwi proteins function during each stage of mouse spermatogenesis has been instrumental in enhancing our understanding of the central mechanisms that underlie piRNA biogenesis and function across mammalian gametogenesis.

1.3.1.1 piRNAs in the fetal testis

Unlike *Drosophila*, the fate of PGCs in mammals is not determined during the initial stages of embryogenesis by the asymmetric segregation of maternal factors (Johnson et al., 2011). Instead, mammalian PGCs differentiate from a population of pluripotent epiblasts via interaction with the surrounding somatic niche later on during embryonic development. As PGCs are derived from the same tissue that make up the rest of the somatic embryo, early embryonic germ cells are not pluripotent and undergo a series of epigenetic transformations to erase the parental DNA methylome contributed by the maternal and paternal germlines. This genome-wide demethylation is followed by a re-establishment of DNA methylation. Fetal testis piRNAs, also called pre-pachytene piRNAs, play a critical role during this time in directing DNA methylation marks to transposable elements, which comprise about half of the mouse genome, to suppress their activity (Aravin et al., 2007, 2008; Gan et al., 2011). During this period of re-establishment, PGCs commit to either a female or male germ cell fate, a decision that is determined by

the sex of the somatic cell niche that interact with early PGCs. PGCs that are induced into a male cell fate differentiate into (pro)spermatogonial cells and undergo a long period of mitotic arrest that lasts until birth. At the same period of time, if PGCs are induced to a female cell fate they will immediately enter meiotic prophase I and then progress through the remaining stages of meiosis in utero (Ewen and Koopman, 2010).

In fetal male mice, the piRNA pathway is first observed in (pro)spermatogonial cells where MILI expression occurs around 12.5 days after fertilization, just before the period in which DNA methylation is re-established at transposon loci (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008). Small RNA sequencing reveals that over half of the piRNAs profiled in the fetal testis are derived from transposable elements (Aravin et al., 2006, 2007). Correspondingly, isolation and sequencing of MILI interacting piRNAs in fetal testis reveal about a third of MILI bound piRNAs map in the sense direction to long interspersed elements (LINEs), short interspersed elements (SINEs), or retrotransposons with long terminal repeats (Aravin et al., 2008). A large minority of MILI bound piRNAs also map to protein coding genes. However, because it is difficult to independently study each of the diverse array of piRNA targets, the specific biological significance of piRNAs mapping to endogenous gene loci during fetal germ cell development is still not clear. Mili mutant males are sterile, and (pro)spermatogonial cells that lack MILI do not progress past meiotic prophase I, which occurs at the time of birth (Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008). These data suggest that the

predominant function of Mili-bound piRNAs is to regulate transposon activity during the critical period in which germ cells are re-establishing DNA methylation.

An emerging feature of the piRNA pathway is the compartmentalization of piRNA biogenesis factors into distinct germ granules with phase separated liquid-liquid-like properties (Ozata et al., 2019). In fetal (pro)spermatogonial cells, MILI interacts with Tudor domain-containing protein TDRD1 at perinuclear granules called pi-bodies (Reuter et al., 2009). In most organisms, the majority of piRNAs are made in phase separated nuage-like granules (Ozata et al., 2019). It is hypothesized that these phase-separated granules act as specialized compartments to promote specific interactions between Piwi proteins and their targets, as well as other piRNA accessory proteins. Furthermore, compartmentalization is thought to exclude other endogenous RNAs from inappropriately interfacing with the piRNA pathway. Supporting this idea, loss of TDRD1 leads to dissolution of pi-bodies and drastically alters the profile of piRNAs that associate with MILI (Reuter et al., 2009; Wang et al., 2009). Like TDRD1, many Tudor domain-containing proteins have been identified to act like scaffolds to sequester piRNA pathway factors to germ granules in other organisms (Aravin et al., 2009; Bortvin, 2013). In addition to Tudor domain-containing proteins, many other piRNA processing factors characterized in the fetal and adult mouse testis are well conserved.

The production of pre-pachetyne piRNAs also requires the endonuclease MITOPLD and the RNA helicase MOV10, homologs of fly Zuc and Armi, respectively (Watanabe et al., 2010). Loss of either MITOPLD or MOV10 leads to defects that are similar to those in mili mutant spermatogonial cells, such as

increase in retrotransposon activity and early meiotic arrest (Aravin et al., 2009). Perinuclear pi-bodies have also been referred to as intermitochondrial cement (IMC), due to the close association between pi-bodies and the surface of mitochondria, where MOV10 and MITOPLD both reside (Aravin et al., 2009; Bortvin, 2013; Shoji et al., 2009). Following MILI expression in early (pro)spermatogonial cells, MIWI2 is expressed from 15 days after fertilization until birth and localizes along with another Tudor domain-containing protein, TDRD9, to a different, larger perinuclear granule, called processing bodies (P-bodies) (Brennecke et al., 2007; Gainetdinov et al., 2018).

MILI and MIWI2 likely interact at these perinuclear germ granules to induce processing of transposon transcripts via ping-pong cycle amplification. Sequencing of MILI and MIWI2 bound piRNAs in the fetal testis has revealed MILI bound piRNAs display prominent features of ping-pong cycle amplification such as a high degree of sequence complementarity to MIWI2 bound piRNAs (Aravin et al., 2008; Shoji et al., 2009). Interestingly, loss of MILI, but not MIWI2, leads to a significant reduction in antisense piRNAs (Manakov et al., 2015), suggesting that while MIWI2-dependent piRNAs can take part in the ping-pong cycle, MILI is the sole driver of homotypic ping-pong cycle amplification. In fly, piRNAs derived from dual-strand clusters initiate piRNA biogenesis (Reuter et al., 2011). In mice, MIWI2-bound piRNAs that are antisense to transposons possess a 10A nucleotide bias (Aravin et al., 2008). Therefore, piRNA biogenesis seems to be initiated from the transposon transcripts instead of piRNA clusters.

Piwi proteins can induce post-transcriptional cleavage or co-transcriptional silencing. Endonucleolytic cleavage ('slicing') of target transcripts in the cytoplasm requires a highly conserved catalytic Asp-Asp-His (DDH) motif in the Piwi proteins (Elbashir et al., 2001a, 2001b; Parker et al., 2004; Schwarz et al., 2004; Yuan et al., 2005). While both MILI and MIWI2 both possess the DDH catalytic triad, only DDH motif mutants in MILI lead to loss of LINE1 (L1) retrotransposon repression and loss of MIWI2 bound piRNAs that map to L1 (Fazio et al., 2011). In contrast, loss of the catalytic DDH triad in MIWI2 does not affect fertility or MIWI2-dependent DNA methylation at L1 elements, although complete loss of MIWI2 leads to severe male sterility (Fazio et al., 2011). These data suggest MIWI2 possesses an alternative function to the silencing of transposon transcripts. Several studies have shown slicing activity is dispensable for Piwi mediated co-transcriptional silencing (Hsieh et al., 2020; Kojima-Kita et al., 2016; Kuramochi-Miyagawa et al., 2008; Vasiliauskaitė et al., 2017; Wang et al., 2015). Instead, Piwi proteins bind to nascent transcripts in the nucleus and regulate gene expression indirectly by recruiting chromatin or transcriptional machinery. Artificial targeting assays, in which MIWI2 is tethered to specific genomic regions that are not normally targeted by the piRNA pathway, reveal that recruitment of MIWI2 is sufficient to induce *de novo* DNA methylation at these regions (Kojima-Kita et al., 2016). Additionally, new evidence reveals that H3K4me2 is transiently enriched at piRNA genomic regions that are targeted for DNA methylation by the piRNA pathway (Nagamori et al., 2018). These H3K4me2 marks are subsequently erased after DNA demethylation and therefore function to specify MIWI2 to these regions for a short window time.

Additionally, the DNA methylase, DNMT3C, is required for de novo DNA methylation at transposons (Barau et al., 2016). However, the molecular mechanism underlying how MIWI2 recruits DNMT3C to transposon-encoding genomic regions is still uncharacterized.

1.3.1.2 Pachytene piRNAs

piRNAs expressed in the adult mouse testis are called pachytene piRNAs because they are specific to the pachytene stage of meiotic prophase I. Pachytene piRNAs are continuously expressed as germ cells cycle through meiosis in the mouse testis starting from 14.5 days after birth through the entire lifespan of the male mouse (Aravin et al., 2006, 2007; Kuramochi-Miyagawa et al., 2004). Pachytene piRNAs are primarily bound by Piwi proteins MIWI, but a significant minority are also bound by MILI. However, unlike MIWI2 and MILI bound pre-pachytene piRNAs, which participate in ping-pong cycle amplification, MILI and MIWI do not show any signatures of ping-pong cycle amplification (Beyret et al., 2012; Vourekas et al., 2012), which suggests that MILI and MIWI likely function independently pachytene staged germ cells.

Small RNA sequencing in *miwi* mutant testis show that MILI-bound piRNAs are upregulated in the absence of MIWI (Fazio et al., 2011). This suggests that MILI and MIWI compete to bind and process the same set of piRNA precursor transcripts. This model is supported by data showing MILI and MIWI bound piRNAs share the same targets, as well as map to the same genomic strand (Vourekas et al., 2012). Additionally, while both MILI and MIWI bound piRNAs show an enrichment for a 5’ U, MILI nor MIWI bound piRNAs seem to display a 10A bias

(Beyret et al., 2012; Vourekas et al., 2012), suggesting ping-pong cycle amplification is no longer active in the adult testis. The main mode of MIWI targeting is post-transcriptional, as research has demonstrated a single site mutation in the DDH catalytic triad of MIWI leads to severe spermiogenesis defects, as well as elevated levels of LINE1 expression (Vourekas et al., 2012), suggesting that maintenance of transposon silencing is still critical in the adult testis. However, methods to sequence RNAs that cross-link directly to MIWI have revealed that these piRNAs that target transposons are a small minority (~20%) (Gou et al., 2014), and that MIWI bound pachytene piRNAs regulate a far more diverse array of transcripts than piRNAs expressed in fetal pre-pachytene germ cells.

The exact functions of pachytene piRNAs are still being actively studied, and several hypotheses have been put forth to explain their biological significance during the pachytene stage. Some evidence suggests piRNAs made during the pachytene stage are required during the later stages of spermatogenesis in post-meiotic, elongating spermatids (Gou et al., 2014; Reuter et al., 2009; Vourekas et al., 2012). Several studies have demonstrated that MIWI coordinates the degradation of a large subset of sperm-specific protein-coding genes (Berthet et al., 2004; Li et al., 2013). Although this regulation is post-transcriptional, it does not require MIWI slicer activity. In elongating spermatids, MIWI targets a subset of protein coding genes for translational repression and mRNA degradation by recruiting the deadenylase CAF1 (Li et al., 2013), a major subunit of testis specific CCR4-CAF1-NOT deadenylase complex. Loss of *caf1* results in severe

spermatogenic defects and sterility similar to miwi mutant spermatids (Zhao et al., 2013). RNA-seq of mRNAs in elongating spermatids derived from *miwi* and *caf1* mutant testes indicate that CAF1 and MIWI specify the same set of target transcripts for degradation (Wu et al., 2020). Following this, piRNA loaded MIWI interacts with the Anaphase-promoting complex (APC) proteasome pathway and undergoes rapid protein degradation in late spermatids (Stein et al., 2015). Interestingly, loss of APC-mediated degradation of MIWI leads to obstruction of sperm maturation, suggesting that MIWI function must be specified for a short window of time.

Only a small subset of piRNAs expressed from a single piRNA generating locus is critical for male fertility (Flemr et al., 2013). Currently, about a hundred mouse loci producing pachytene piRNAs have been identified, and all require the male-specific transcription factor, A-MYB, for their transcription (Li et al., 2013). Interestingly, disrupting each piRNA locus leads to distinct outcomes. For example, loss of a single piRNA locus encoded on chromosome 17 results in the loss of 17% of pachytene piRNAs, but seems to have no apparent effect on testis development (Homolka et al., 2015). In contrast, loss of another piRNA locus, pi6, which generates only 5.8% of all pachytene piRNAs, leads to severe male-specific sterility (Flemr et al., 2013). Study of pi6 mutants show that they can still can repress transposons, but are defective in fertilization and display elevated levels of mRNAs from genes involved in sperm penetration of the oocyte zona pellucida, an outer glycoprotein layer that surrounds mammalian oocytes that sperm must penetrate for successful fertilization (Hsieh et al., 2020). Consistent with this

finding, pi6 mutant sperm are defective in their ability to fertilize oocytes in vitro, and are unable to penetrate the zona pellucida as effectively as wild-type sperm (Malki et al., 2014). A small number of *pi6* mutant sperm are able to fertilize oocytes. However, heterozygous *pi6* mutant embryos abort gestation at high frequencies compared to wild-type embryos. The *pi6* locus targets a very small number of genes with biological significance. Instead, piRNAs transcribed from the *pi6* locus map to four other piRNA expressing loci, in *trans* (Roovers et al., 2015; Tan et al., 2020) and aid in ramping up the production of piRNAs from these four downstream piRNA clusters. Additionally, this suggests that each pachytene piRNA cluster have evolved independent targets and biological functions.

1.3.1.3 The role of piRNAs during meiotic chromosome segregation

Recent research has uncovered additional, novel functions for the piRNA pathway in regulating the proper segregation of chromosomes during meiosis. Preliminary research suggests Miwi, alongside Dicer-dependent siRNAs, localizes to the nucleus to prevent the mis-segregation of chromosomes by targeting major and minor satellite transcripts from large centromeric repeats (Camacho et al., 2017; Chan et al., 2012). Interestingly, MILI is not needed for the suppression of satellite RNA, however MILI interacting piRNAs that map to satellite RNAs are lost when MIWI's slicing activity is inhibited (Williams et al., 2015). The exact function of these non-coding satellite RNAs is not well characterized, however, emerging evidence reveals that they are key players in diverse processes that are required for appropriate meiotic segregation of chromosomes such as proper kinetochore

assembly, maintenance of pericentric heterochromatin, and the establishment of centromere identity (Gu et al., 2012).

1.3.2 piRNA function during murine oogenesis

Although several studies show piRNAs are present in fetal oocytes, the piRNA pathway does not appear to play an active role during mouse fetal oocyte development. Curiously, while LINE1 is genotoxic and therefore silenced by piRNAs in male germ cells, LINE1 elements are active in fetal oocytes and are part of a conserved process called fetal oocyte attrition (FOA) (Ellis, 2006) Before entering meiotic prophase I, about 80% of fetal oocytes are eliminated by FOA. LINE1 elements are the primary contributor of FOA through DNA damage-driven apoptosis (Malki et al., 2014).

Closer inspection of LINE1 activity across a population of fetal oocytes before FOA reveal that oocytes possess a wide variety of LINE1 activity, with a minor fraction of oocytes expressing little to no LINE1 transcripts (Malki et al., 2014). Oocytes with low LINE1 expression, are observed to circumvent FOA and progress to meiotic prophase I. Interestingly, while the loss of MIWI, MIWI2, or MILI has no effect on FOA, the loss of the conserved piRNA pathway protein, Maelstrom (MAEL), leads to an increase in LINE1 activity, enhanced meiotic defects, as well as significantly higher frequencies of FOA (Malki et al., 2014). piRNAs are hypothesized to act as sequence-specific guides to recruit nuclear factors like MAEL to genomic transposon loci. Therefore, how MAEL can target transposon species in fetal oocytes without a piRNA guide remains unclear.

Endogenous small interfering RNAs (endo-siRNAs) may play a redundant function with piRNAs to direct silencing machinery, such as MAEL, to transposon loci (Consortium*, 1998; Orgel and Crick, 1980). However, female mice that do not express DICER, required for the production of the majority of endo-siRNAs in oocytes, or MILI are still fertile, suggesting alternative mechanisms are required for limiting LINE1 expression (Flemr et al., 2013; Murchison et al., 2007). Nonetheless, these data suggest that LINE1 activity is used as a gauge for the fetal female germline to select oocytes with a healthy reproductive capacity with minimal deleterious changes to the genome. Compared to spermatogenesis, oogenesis is a far more energy and nutrient consuming process. Therefore, the female germline may be more inclined to sacrifice oocytes with active transposons early on during development, rather than expend energy to maintain an active piRNA silencing pathway.

1.3.3 piRNAs in non-murine mammalian systems

Recent characterization of Piwi proteins during human and bovine gametogenesis show many features of the piRNA pathway in mice may not be conserved in other mammalian species. For example, the localization of mouse and human Piwi proteins differ substantially. MIWI2 is enriched in perinuclear P-bodies, while its human ortholog, PIWIL4, localizes to IMC like germ granules that associate closely with the outer mitochondrial membrane (Fernandes et al., 2017). Furthermore, mouse MILI localizes to P-bodies, while its human ortholog PIWIL2, localizes to small granules that are scattered throughout the cytoplasm instead of showing strong association with PIWIL4 containing germ granules. These data

suggest that unlike MIWI2 and MILI which closely interact in the mouse fetal testis, PIWIL2 and PIWIL4 act separately. However, few genetic or biochemical studies have been performed to further elucidate their interaction in non-murine mammalian systems.

Sequencing of human piRNAs in different developmental stages of oogenesis suggests that fetal ovaries harbor a distinct piRNA repertoire compared to adult ovaries (Shen et al., 2018a). The human genome encodes a fourth Piwi protein, PIWIL3, which is not found in the mouse genome. Study of PIWIL3 colocalization in bovine fetal oocytes, suggests PIWIL3 may interact with conserved Tudor domain protein, TDRKH, responsible for pre-piRNA 3' end trimming, as well as a 3' exonuclease, PNLDC1, required for processing of piRNAs (Bagijn et al., 2012; Shen et al., 2018b, 2018a; Zhang et al., 2018). In addition, the vast majority of PIWIL3 bound piRNAs map antisense to transposons, suggesting they play a similar role to MILI bound fetal piRNAs in repressing transposons in mouse gonocytes. However, the consequences of PIWIL3 mutations are not known and, therefore, the biological significance of PIWIL3-bound piRNAs is still not completely understood.

1.4 piRNAs across *C. elegans* germline development

C. elegans are hermaphrodites that undergo spermatogenesis during late larval stages, then switch to oogenesis during adulthood (Lee et al., 2012). In addition, *C. elegans* are transparent and all stages of germ cell development are represented within a single adult animal, making them a powerful genetic system to visualize piRNA dynamics across sexual development. Over 15,000 piRNAs are

densely encoded within two mega-base genomic clusters on chromosome IV (Gu et al., 2012; Ruby et al., 2006). Unlike mature piRNAs from other species, *C. elegans* piRNAs are not processed from long multi-kilobase precursor transcripts. Instead, each *C. elegans* piRNA is individually transcribed from a discrete locus and produces 26-29 nt precursors (Gu et al., 2012; Ruby et al., 2006). Precursor processing gives rise to mature piRNAs that are called 21U RNAs due to their 5' uracil and 21 nt length. 21U RNAs are loaded into a conserved Piwi protein, PRG-1, which has been shown to bind to a large range of target transcripts (Lee et al., 2012; Shen et al., 2018a).

Emerging evidence suggests that the piRNA pathway is involved not only in silencing repetitive elements, but also in diverse biological processes. Most genomes are replete with transposons (about half the mouse genome encodes repetitive elements such as transposons). Therefore, developing genetic tools to study the piRNA pathway in the absence of active transposons has been a challenge. In contrast, only 12% of the *C. elegans* genome is derived from transposons, and most are no longer mobile, thus making it an ideal system to study alternative piRNA functions. While a small minority of 21U RNAs do silence transposons (Das et al., 2008), the vast majority of 21U RNAs are thought to target protein coding gene transcripts (Batista et al., 2008; Lee et al., 2012; Shen et al., 2018a; Shirayama et al., 2012).

Analogous to microRNAs, PRG-1 loaded 21U RNAs tolerate mismatches when they bind to their complementary targets. A method capable of identifying direct piRNA-target hybrids has recently demonstrated that PRG-1 binds to

transcripts of almost all protein coding genes expressed during hermaphrodite oogenesis (Shen et al., 2018a). Although PRG-1 contains the conserved DDH catalytic motif and is therefore capable of slicing RNA targets in vitro (Bagijn et al., 2012), PRG-1 does not generally slice it's target transcripts.

1.4.1 piRNA function during *C. elegans* oogenesis

The ping-pong piRNA amplification cycle is not present in nematode species. Instead, the *C. elegans* piRNA pathway mainly acts upstream of another class of endogenous small interfering RNAs, the 22G RNAs, to silence target transcripts (Ashe et al., 2012; Bagijn et al., 2012; Das et al., 2008). The 22G RNAs display a 5' guanosine bias and are 22 nt in length (Ambros et al., 2003; Collins et al., 1987) and are made by RNA-dependent RNA polymerases (RdRPs) that use mRNAs as direct templates for their synthesis (Aoki et al., 2007; Pak and Fire, 2007; Sijen et al., 2007). PRG-1 binding triggers the recruitment of RdRPs to transcripts resulting in 22G RNA production. 22G RNAs then interact with members of a large, 12-member class of worm Argonaute proteins (WAGOs) (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2009; Guang et al., 2008; Ni et al., 2014; Shirayama et al., 2012; Sijen et al., 2007). Studies in *C. elegans* have revealed that 21U RNA pathway is required for WAGO 22G RNA dependent transmission of epigenetic information over multiple generations. However, to what extent each WAGO protein functions under the 21U RNA pathway is not completely understood, though the characterization of each WAGO is actively underway (reviewed in Weiser and Kim 2019).

Studies in *C. elegans* have been pivotal in elucidating the critical role that endogenous small RNAs, including 21U RNAs, play in transgenerational inheritance. The most well-characterized player is HRDE-1, a nuclear localized WAGO protein that acts downstream of PRG-1 to maintain the multigenerational silencing initiated by PRG-1 in the maternal germline. Through downstream WAGO-22G RNAs that are transmitted through the maternal germline, PRG-1 can silence foreign transgenes which are paternally inherited (Luteijn et al., 2012). This knowledge has been proven useful for expressing transgenes in the germline, as transgenes that are silenced by PRG-1 can be engineered for higher expression in the maternal germline by changing sequences heavily targeted by the 21U RNA pathway (Zhang et al., 2018). Recent analysis reveals that WAGO 22G RNAs are able to sustain this silencing, without 21U RNAs, as loss of maternal PRG-1 still results in robust silencing of a GFP transgene. However, maternal re-introduction of PRG-1 results in misrouting of WAGO 22G RNA to other endogenous gene targets as well as causing substantial loss in fertility (Phillips et al., 2015). This suggests that PRG-1 acts as a powerful form of surveillance in the female germline, to ensure that transcripts are specified for 22G RNA mediated silencing.

How does the germline ensure that germline genes are expressed, if the majority of protein coding transcripts in the germline are bound by PRG-1? Evidence from several studies suggest that another WAGO protein, CSR-1, bind nascent transcripts in the nucleus to antagonize PRG-1 function and promote the stability of endogenous germline transcripts (Shirayama et al., 2012). Analysis of transcriptome-wide PRG-1 and CSR-1 22G RNA binding profiles reveals that the

binding density of PRG-1 at protein coding genes is anticorrelated with CSR-1 Additionally, loss of CSR-1 function in the female germline is not observed to de-silence a significant portion of germline genes, instead, a moderate subset of genes is slightly downregulated (Claycomb et al., 2009). These data suggest CSR-1 plays an activating role at protein coding genes during oogenesis by protecting against PRG-1 surveillance. In addition to surveillance of potential foreign elements in the germline, new evidence reveals the 21U RNA pathway may directly repress endogenous protein coding genes, independent of CSR-1, that play critical roles during germline development and differentiation (Tang et al., 2018).

One gene in particular, *xol-1*, plays an important role in X chromosome dosage compensation as well as sex determination in *C. elegans* (Tang et al., 2018). Study of hermaphrodites as well as its progeny that have lost a single 21U-RNA, *21ux-1*, that targets *xol-1* show 21U RNAs, inherited maternally, function to promote dosage compensation in the zygotes of the next generation by sensitizing them to X-signal elements (Tang et al., 2018). The loss of *21ux-1* leads to an increase in XOL-1 expression compared to wildtype hermaphrodites. Accordingly, *xol-1*, which is normally suppressed and dispensable for hermaphrodite sexual development, but critical for males, becomes upregulated in the absence of 21ux-1. Further loss of X-signal elements, like *sex-1*, which acts redundantly with 21ux-1 to repress *xol-1*, lead to a higher frequency in the masculinization of hermaphrodites as well as other phenotypes indicative of defective dosage compensation (Tang et al., 2018).

1.4.2 piRNA function during *C. elegans* spermatogenesis

C. elegans undergoes spermatogenesis during late larval stages before switching to oogenesis during early adulthood. Spermatogenic 21U RNAs make up over half the 21U RNAs expressed in the germline (Billi et al., 2013; Choi et al., 2020), however as discussed previously, only maternal, and not paternal 21U RNAs, are transmitted to the next generation or play an active role in 22G RNA mediated transgenerational inheritance. Loss of PRG-1 function in males, leads to severe loss in fertility as well as the downregulation of a moderate number of sperm specific genes (Wang and Reinke, 2008), suggesting PRG-1 plays a role in promoting the expression of genes important for spermatogenesis.

1.5 Sex-specific piRNAs are defined at the transcriptional level

The transcription of piRNAs in different organisms show multiple instances of divergence from canonical transcription at the level of initiation, elongation, and termination. In *Drosophila*, the transcription of dual-strand clusters is driven by chromatin factors that bind heterochromatin regions normally associated with transcriptional quiescence. In *C. elegans*, the germline has co-opted the use of different snRNA transcription machinery subunits to drive sex-specific piRNA transcription. In mammals, pachytene stage piRNAs require an ancient male germline specific transcription factor, A-Myb, essential for the transcription of other genes involved in regulating meiotic cell cycle control specifically during mammalian spermatogenesis.

1.5.1 Promoting piRNA transcription at heterochromatin regions in *Drosophila*

In the *Drosophila* germline, the majority of piRNAs are expressed from large dual-strand clusters that must license transcription from heterochromatin regions. How does transcriptional machinery distinguish piRNA transcriptional loci from other heterochromatic loci? Dual-strand clusters are a unique example of co-option, where the piRNA pathway promotes transcription at heterochromatin, normally harboring gene-poor and transcriptionally quiescent genomic regions. piRNA transcription is directed by, Rhino, an HP1 homolog that directly binds H3K9me3 marks at dual-strand clusters via its chromodomain to recruit additional piRNA transcription factors, Cutoff and Deadlock, to form the Rhino-Deadlock-Cutoff (RDC) complex (Andersen et al., 2017; Chen et al., 2016; Klattenhoff et al., 2009; Mohn et al., 2014; Pane et al., 2011). How Rhino is able to distinguish other regions of heterochromatin from dual-strand clusters is still under active investigation.

In fly oocytes, Rhino, Deadlock, as well as Cutoff must interact with H3K9me3 marks in heterochromatic regions to drive piRNA precursor transcription from dual-strand piRNA clusters. Due to positive selection, Rhino, Deadlock, and Cutoff are evolving at a rapid pace. For example, cross species studies between *D. melanogaster* and sibling species *Drosophila simulans* show Rhino, Deadlock, or Cutoff orthologs are not interchangeable in function (Parhad et al., 2017). D. simulans Cutoff acts as a dominant negative allele, when expressed as a transgene, and cannot rescue the loss of piRNA transcription in *D. melanogaster*

cutoff mutants (Parhad et al., 2017). The functional domains of RDC members are evolving at a rapid pace, which is a hallmark of adaptive evolution highlighting the arms race between germline and rapidly evolving toxic genetic elements.

1.5.2 Transcriptional regulation of mammalian pachytene piRNAs

Unlike the transcriptional regulation of pre-pachytene piRNAs, which remains elusive, new emerging evidence brings insight into how pachytene piRNAs are transcriptionally regulated, and demonstrates that transcriptional regulation in mammals is highly sex-specific. For example, A-Myb, a master regulator of male meiosis, as it is critical for the transcription of genes that regulate the meiotic cell cycle (Gu et al., 2012), is also required for the production of pachytene piRNAs. A-Myb dependent piRNAs are transcribed from about one hundred well defined pachytene piRNA loci that encompass the majority of piRNAs expressed during pachytene stages germ cell development (Weick et al., 2014). In addition to coordinating the transcription of pachytene piRNA precursors as well as meiotic genes, A-Myb binding sites are found within its own promoter and the promoters of genes encoding piRNA biogenesis factors, including Miwi, MitoPld, and Tdrd9. A-Myb is shown to be first expressed as spermatogonial cells enter meiosis I, and additional binding of piRNA biogenesis factors by A-Myb is thought to ramp up the production of piRNA biogenesis machinery to facilitate the large flux in piRNA precursor transcription that occurs as germ cells enter meiosis.

The promoters of piRNA loci as well as the processing of nascent piRNA precursor transcripts are atypical but also resemble features common to genes licensed specifically during spermatogenesis. Unlike somatic genes, pachytene

piRNA clusters are depleted of CG dinucleotides that are usually methylated and associated with transcriptional quiescence in the soma. Additionally, pachytene piRNA clusters are enriched with H3K4me3 and H3K9ac (Hammoud et al., 2014; Li et al., 2013), which are hypothesized to promote transcription.

1.5.3 piRNA biogenesis in *C. elegans*

In *C. elegans*, the vast majority of the ~15,000 21U RNAs that are transcribed from two large clusters on chromosome IV, requires the joint coordination of a diverse array of germline nuclear factors, some of which have evolved from alternate transcriptional processes such as snRNA biogenesis. Such coordination forms distinct nuclear foci that mark the chromosome location of these clusters in pre-pachytene and pachytene germ cell nuclei.

Transcription of 21U RNAs in species of the Rhabiditina clade requires a highly conserved upstream 8-nt sequence, CTGTTTCA, called the Ruby motif, which is separated from each 21U RNA encoding locus by an AT rich spacer ranging from 25-60 nt in length (Billi et al., 2013). While the exact mechanism by which the Ruby motif drives 21U RNA transcription is still not understood, it is hypothesized the Ruby motif acts as a transcription factor binding site. Recent evolutionary analysis of piRNA promoters across different nematode species show that the Ruby motif originated from snRNA promoter cis-elements (Beltran et al., 2019). SNPC-4, the largest DNA binding subunit of the small nuclear RNA activating complex (SNAPc), is critical for the transcription of the vast majority of 21U RNAs expressed during germline development (Choi et al., 2021; Kasper et al., 2014). In mammalian cells, SNAPC4, the homolog of the *C. elegans* SNAPC-

4, must form a complex with two additional proteins SNAPC1 and SNAPC3 in order to bind the proximal sequence element (PSE) of snRNA loci to drive snRNA transcription in vitro. This suggests that SNPC-4 is an evolutionary conserved transcription factor that interacts with the Ruby motif. However, SNPC-4 is also required for the expression of 21U RNAs of a less abundant 21U RNA species that is proximal to protein coding genes that do not contain a Ruby motif.

While SNPC-4 is required for both types, a recent study found PRDE-1, a casein kinase I family protein specifically enriched in the germline, recruits SNPC-4 to the Ruby motif to drive 21U RNA transcription (Sarkies et al., 2015). PRDE-1 interacts with SNPC-4 and loss of PRDE-1 leads to the dissolution of SNPC-4 foci at the chromosome IV clusters, wheras 21U RNA expression from protein coding loci remains unperturbed. These data suggest that PRDE-1 may interact specifically with the Ruby motif and recruit SNPC-4 to 21U RNA loci. A genome-wide RNAi-based screen to identify piRNA biogenesis factors has recently identified three additional 21U RNA transcription factors, Twenty One U-RNA biogenesis Fouled Up 3, 4, and 5 (TOFU-3,4 and 5). Recent evidence shows TOFU-4 and TOFU-5 are also required for 21U RNA biogenesis and colocalize with SNPC-4 and PRDE-1 containing foci (Goh et al., 2014); however further study is required to understand the molecular role that TOFU-4 and TOFU-5 play at these foci.

In *C. elegans*, 21U RNAs peak in expression during the fourth larval stage, where they undergo spermatogenesis, as well as during adulthood when the germline undergoes oogenesis. Although the 21U RNA pathway is active across

both male and female germline development, the majority 21U RNAs (~70%) are differentially expressed in male versus female germline (Billi et al., 2013; Choi et al., 2021). Studies using a transgenic system demonstrates the conserved Ruby motif is critical for regulating this sex specificity, as the identity of the 5' nucleotide of the Ruby motif directs sex-biased 21U RNA expression across hermpahrodite germline development. The 5' cytidine is required for male specificity because mutation of the cytidine to an adenosine, within a transgene containing a male 21U RNA locus, leads to male as well as female germline expression (Billi et al., 2013). In contrast, replacement of a 5' adenosine of the Ruby motif within a transgene encoding a female locus, to a cytidine, leads to complete loss of expression in both sexes (Billi et al., 2013). This demonstrates that the 5' nucleotide identity is critical for sex-specific transcriptional regulation; however, identification of trans-acting factors that interact specifically with the Ruby motif *in vivo*, let alone direct Ruby motif-dependent sex-biased 21U RNA expression has been difficult. This may in part be due to the high density of 21U RNA expressing loci, making it difficult to deduce the accurate transcriptional machinery footprint at each 21U RNA loci on a genome-wide scale. Methods such as Chip-Exo (Skene and Henikoff, 2015) have the potential of improving the binding resolution of these factors and may provide sharper insight into the interaction between trans-factors and cis-motifs that govern 21U RNA transcription.

The study of evolution of the piRNA pathway within nematodes highlights one extreme end of the diversity that governs function of piRNAs. Strikingly, the piRNA pathway is not present in most nematode lineages outside of Rhabiditina.

Closer inspection of nematode species lacking a piRNA pathway reveal they have retained a more ancestral system, using RdRP generated small RNAs that are also prevalent in plants and fungi, to combat transposable elements (Sarkies et al., 2015). In contrast, mammalian and insect systems do not encode for RdRPs, and it is hypothesized the piRNA pathway has replaced the ancestral RdRP pathway, to counter foreign elements. The regulation of 21U RNA biogenesis in *C. elegans* by snRNA transcription factors, such as SNPC-4, is likely an even more recent adaptation.

1.6 Concluding Remarks

In summary, piRNA expression is regulated by a rapidly evolving network of pIRNA biogenesis machinery, whereas the ancestral effector functions of Piwi proteins remain conserved. Since the initial discovery of piRNAs 17 years ago, several fundamental questions still remain. For example, Rhino dependent dual-strand cluster piRNAs are specific to *D. melanogaster*. While dual-strand clusters are present in other drosophilids (Parhad et al., 2017), Rhino homologs have not been identified outside of *D. melanogaster*. Studies into what other factors have evolved to drive piRNA expression from these clusters in other insect species is still unknown and could provide further insight into the evolution of piRNA biogenesis.

A growing body of evidence suggests that the diversity of piRNA function across different organisms and between sexes is likely due to the rapidly evolving *cis*-regulatory elements and chromatin landscape at each piRNA loci, such that piRNA loci are extremely distinct even between closely related sister species.

Despite this diversity in piRNA sequence and function, many piRNA biogenesis factors downstream of transcription are well conserved in function. For example, piRNA processing factors like Zucchini and Armitage, as discussed previously, show conserved function in the production of phased piRNAs in both insect and mammalian systems.

A recent study performed a meta-analysis of small RNA sequencing datasets derived from 33 different species, from nematode to human, and identified features that are common across different organisms harboring an active piRNA pathway. Findings from this analysis show signatures of ping-pong mediated amplification as well as phased piRNA production, mediated by Zucchini in mouse and fly, are conserved in 32 of the 33 species assessed, suggesting the machinery underlying most piRNA processing downstream of transcription have existed for at least 800 million years (Cecere et al., 2012).

Nematode piRNA processing is the only exception, as each *C. elegans* piRNA-precursor only produces a single mature piRNA. Despite these differences, the *C. elegans* piRNA pathway still carries out its ancestral functions in silencing transposon elements, specifically a family of DNA transposons, via the conserved Piwi family of protein argonautes.

In contrast to downstream piRNA processing machinery, piRNA transcription factors display remarkable phylogenetic diversity. Many such factors seem to have evolved rapidly under positive selection, a hallmark of a host-pathogen arms race. Transposons are an ancient genetic pathogen that have shaped many aspects of metazoan evolution. For a transposon species to survive,

it must continuously evade host detection driving a rapid change in transposon sequence. In turn, the piRNA pathway must continuously update its sequence repertoire to accurately target constantly evolving transposon species. This arms race between transposons and the piRNA pathway has resulted in diverse piRNA biogenesis mechanisms, especially at the level of transcription.

1.7 References

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and Other Tiny Endogenous RNAs in C. elegans. Current Biology *13*, 807–818.

Andersen, P.R., Tirian, L., Vunjak, M., and Brennecke, J. (2017). A heterochromatin-dependent transcription machinery drives piRNA expression. Nature *549*, 54–59.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. EMBO Rep. *26*, 5007–5019.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., Klenov, M.S., Vagin, V.V., Bantignies, F., Cavalli, G., and Gvozdev, V.A. (2004). Dissection of a natural RNA silencing process in the *Drosophila melanogaster* germ line. Mol. Cell. Biol. *24*, 6742–6750.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744–747.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell 31, 785–799.

Aravin, A.A., Heijden, G.W. van der, Castañeda, J., Vagin, V.V., Hannon, G.J., and Bortvin, A. (2009). Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. PLoS Genet. 5, e*10*00764.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (20*12*). piRNAs Can Trigger a Multigenerational Epigenetic Memory in the Germline of C. elegans. Cell *150*, 88–99.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E.A. (2012). Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. Science *337*, 574–578.

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science *354*, 909–912.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and

21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. Mol. Cell 31, 67–78.

Beltran, T., Barroso, C., Birkle, T.Y., Stevens, L., Schwartz, H.T., Sternberg, P.W., Fradin, H., Gunsalus, K., Piano, F., Sharma, G., et al. (2019). Comparative epigenomics reveals that RNA polymerase II pausing and chromatin domain organization control nematode piRNA Biogenesis. Dev Cell *48*, 793-810.

Berkseth, M., Ikegami, K., Arur, S., Lieb, J.D., and Zarkower, D. (2013). TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination. Proc. Nat. Acad. Sci. *110*, 16033–16038.

Beyret, E., Liu, N., and Lin, H. (2012). piRNA biogenesis during adult spermatogenesis in mice is independent of the ping-pong mechanism. Cell Res *22*, 1429–1439.Berthet, C., Morera, A.-M., Asensio, M.-J., Chauvin, M.-A., Morel, A.-P., Dijoud, F., Magaud, J.-P., Durand, P., and Rouault, J.-P. (2004). CCR4-Associated factor CAF1 is an essential factor for spermatogenesis. Mol. Cell Biol. *24*, 5808–5820.

Billi, A.C., Freeberg, M.A., Day, A.M., Chun, S.Y., Khivansara, V., and Kim, J.K. (2013). A conserved upstream motif orchestrates autonomous, germline-enriched expression of *Caenorhabditis elegans* piRNAs. PLoS Genet. *9*, e1003392.

Bolcun-Filas, E., Bannister, L.A., Barash, A., Schimenti, K.J., Hartford, S.A., Eppig, J.J., Handel, M.A., Shen, L., and Schimenti, J.C. (2011). A-MYB (MYBL1)

transcription factor is a master regulator of male meiosis. Dev *138*, 3319–3330.

Bono, M. de, and Hodgkin, J. (1996). Evolution of sex determination in *Caenorhabditis*: unusually high divergence of *tra-1* and its functional consequences. Genetics *144*, 587–595.

Bortvin, A. (2013). PIWI-interacting RNAs (piRNAs) — a mouse testis perspective. Biochem. Mosc. *78*, 592–602.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089–1103.

Brower-Toland, B., Findley, S.D., Jiang, L., Liu, L., Yin, H., Dus, M., Zhou, P., Elgin, S.C.R., and Lin, H. (2007). *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. Gene Dev *21*, 2300–2311.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447–451.

Camacho, O.V., Galan, C., Swist-Rosowska, K., Ching, R., Gamalinda, M., Karabiber, F., Rosa-Velazquez, I.D.L., Engist, B., Koschorz, B., Shukeir, N., et al. (2017). Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. eLife *6*, e25*29*3.

Carmell, M.A., Girard, A., Kant, H.J.G. van de, Bourc'his, D., Bestor, T.H., Rooij, D.G. de, and Hannon, G.J. (2007). MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. Dev Cell *12*, 503–514.

Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012). Promoters recognized by Forkhead proteins exist for individual 21U-RNAs. Mol. Cell *47*, 734–745.

Chan, F.L., Marshall, O.J., Saffery, R., Kim, B.W., Earle, E., Choo, K.H.A., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc. Natl. Acad. Sci. *109*, 1979–1984.

Chen, P., and Ellis, R.E. (2000). TRA-1A regulates transcription of *fog-3*, which controls germ cell fate in *C. elegans*. Dev. *127*, 3119–3129.

Chen, P., Luo, Y., and Aravin, A.A. (2020). RDC complex executes a dynamic piRNA program during *Drosophila* spermatogenesis to safeguard male fertility. bioRxiv.

Choi, C.P., Tay, R.J., Starostik, M.R., Feng, S., Moresco, J.J., Montgomery, B.E., Xu, E., Hammonds, M.A., Schatz, M.C., Montgomery, T.A., et al. (2021). SNPC-1.3 is a sex-specific transcription factor that drives male piRNA expression in *C. elegans*. eLife *10*, e60681.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., Wolfswinkel, J.C. van, Chaves, D.A., Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009).

The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell *139*, 123–134.

Collins, J., Saari, B., and Anderson, P. (1987). Activation of a transposable element in the germ line but not the soma of Caenorhabditis elegans. Nature *328*, 726–728.

Consortium*, T. *C. elegans* S. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science *282*, 2012–2018.

Cox, D.N., Chao, A., and Lin, H. (2000). *piwi* encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Dev. *127*, 503–514.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. Mol. Cell *31*, 79–90.

Deng, W., and Lin, H. (2002). miwi, a Murine Homolog of piwi, Encodes a Cytoplasmic Protein Essential for Spermatogenesis. Dev Cell *2*, 819–830.

Doniach, T., and Hodgkin, J. (1984). A sex-determining gene, *fem-1*, required for both male and hermaphrodite development in *Caenorhabditis elegans*. Dev. Biol *106*, 223–235.

Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. (2001a). Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. Embo J *20*, 6877–6888.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Gene Dev *15*, 188–200.

Ellis, R. (2006). Sex determination in the germ line. In Wormbook, The *C. elegans* Research Community, WormBook ed., doi/10.1895/wormbook.1.82.2, http://www.wormbook.org.

Ewen, K.A., and Koopman, P. (2010). Mouse germ cell development: From specification to sex determination. *Mol Cell Endocrinol* 323, 76–93.

Fazio, S.D., Bartonicek, N., Giacomo, M.D., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P.N., Enright, A.J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* 480, 259.

Fernandes, M.G., He, N., Wang, F., Iperen, L.V., Eguizabal, C., Matorras, R., Roelen, B.A.J., and Lopes, S.M.C.D.S. (2017). Human-specific subcellular compartmentalization of P-element induced wimpy testis-like (PIWIL) granules during germ cell development and spermatogenesis. Hum. Reprod. *33*, 258–269.

Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven Dicer isoform directs

endogenous small interfering RNA production in mouse oocytes. Cell *155*, 807–816.

Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K., and Zamore, P.D. (2018). A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. Mol. Cell *71*, 775-790.

Gan, H., Lin, X., Zhang, Z., Zhang, W., Liao, S., Wang, L., and Han, C. (2011). piRNA profiling during specific stages of mouse spermatogenesis. RNA *17*, 1191–1203.

Goh, W.-S.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. Genes Dev. *28*, 797–807.

Gonzalez, J., Qi, H., Liu, N., and Lin, H. (2015). Piwi Is a key regulator of both somatic and germline stem cells in the *Drosophila* testis. Cell Reports *12*, 150–161.

Gou, L.-T., Dai, P., Yang, J.-H., Xue, Y., Hu, Y.-P., Zhou, Y., Kang, J.-Y., Wang, X., Li, H., Hua, M.-M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res *24*, 680–700.

Goriaux, C., Desset, S., Renaud, Y., Vaury, C., and Brasset, E. (2014). Transcriptional properties and splicing of the *flamenco* piRNA cluster. EMBO Rep. *15*, 411–418.

Gou, L.-T., Dai, P., Yang, J.-H., Xue, Y., Hu, Y.-P., Zhou, Y., Kang, J.-Y., Wang, X., Li, H., Hua, M.-M., et al. (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. *24*, 680–700.

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. Mol Cell *36*, 231–244.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. Cell *151*, 1488–1500.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute Transports siRNAs from the Cytoplasm to the Nucleus. Science *321*, 537–541.

Hammoud, S.S., Low, D.H.P., Yi, C., Carrell, D.T., Guccione, E., and Cairns, B.R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. Cell Stem Cell *15*, 239–253.

Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science *348*, 817–821.

Homolka, D., Pandey, R.R., Goriaux, C., Brasset, E., Vaury, C., Sachidanandam, R., Fauvarque, M.-O., and Pillai, R.S. (2015). PIWI Slicing and RNA Elements in Precursors Instruct Directional Primary piRNA Biogenesis. Cell Reports 12, 418–428.

Hsieh, C., Xia, J., and Lin, H. (2020). MIWI prevents aneuploidy during meiosis by cleaving excess satellite RNA. Embo J 39, e103614.

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature *491*, 279–283.

Johnson, A.D., Richardson, E., Bachvarova, R.F., and Crother, B.I. (2011). Evolution of the germ line–soma relationship in vertebrate embryos. Reproduction *141*, 291–300.

Kasper, D.M., Wang, G., Gardner, K.E., Johnstone, T.G., and Reinke, V. (2014). The *C. elegans* SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell *31*, 145–158.

Kiuchi, T., Koga, H., Kawamoto, M., Shoji, K., Sakai, H., Arai, Y., Ishihara, G., Kawaoka, S., Sugano, S., Shimada, T., et al. (2014). A single female-specific piRNA is the primary determiner of sex in the silkworm. Nature *509*, 633–636.

Klattenhoff, C., Bratu, D.P., McGinnis-Schultz, N., Koppetsch, B.S., Cook, H.A., and Theurkauf, W.E. (2007). *Drosophila* rasiRNA pathway mutations

disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. Dev. Cell *12*, 45–55.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The *Drosophila* HP1 homolog Rhino Is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137–1149.

Kojima-Kita, K., Kuramochi-Miyagawa, S., Nagamori, I., Ogonuki, N., Ogura, A., Hasuwa, H., Akazawa, T., Inoue, N., and Nakano, T. (2016). MIWI2 as an Effector of DNA Methylation and Gene Silencing in Embryonic Male Germ Cells. Cell Reports 16, 2819–2828.

Kotelnikov, R.N., Klenov, M.S., Rozovsky, Y.M., Olenina, L.V., Kibanov, M.V., and Gvozdev, V.A. (2009). Peculiarities of piRNA-mediated post-transcriptional silencing of *Stellate* repeats in testes of *Drosophila melanogaster*. Nucleic Acids Res. *37*, 3254–3263.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., et al. (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839–849.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Gene Dev. *22*, 908–917.

Lamont, L.B., and Kimble, J. (2007). Developmental expression of FOG-1/CPEB protein and its control in the *Caenorhabditis elegans* hermaphrodite germ line. Dev. Dynam. *236*, 871–879.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). *C. elegans* piRNAs mediate the genome-wide surveillance of germline transcripts. Cell *150*, 78–87.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., et al. (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509–521.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., et al. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol. Cell *50*, 67–81.

Lim, A.K., and Kai, T. (2007). Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. *104*, 6714–6719.

Luteijn, M.J., Bergeijk, P. van, Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in *Caenorhabditis elegans*. EMBO *31*, 3422–3430.

Malki, S., van der Heijden, G.W., O'Donnell, K.A., Martin, S.L., and Bortvin, A. (2014). A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. Dev. Cell *29*, 521–533.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA Pathways Act in Germline and Somatic Tissues of the *Drosophila* Ovary. Cell 137, 522–535.

Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The Rhino-Deadlock-Cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. Cell *157*, 1364–1379.

Mohn, F., Handler, D., and Brennecke, J. (2015). piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. Science *348*, 812–817.

Moshkovich, N., and Lei, E.P. (2010). HP1 Recruitment in the Absence of Argonaute Proteins in *Drosophila*. Plos Genet *6*, e1000880.

Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. Gene Dev *21*, 682–693.

Nagamori, I., Kobayashi, H., Nishimura, T., Yamagishi, R., Katahira, J., Kuramochi-Miyagawa, S., Kono, T., and Nakano, T. (2018). Relationship

between PIWIL4-Mediated H3K4me2 Demethylation and piRNA-Dependent DNA Methylation. Cell Reports *25*, 350–356.

Nagao, A., Mituyama, T., Huang, H., Chen, D., Siomi, M.C., and Siomi, H. (2010). Biogenesis pathways of piRNAs loaded onto AGO3 in the *Drosophila* testis. RNA *16*, 2503–2515.

Ni, J.Z., Chen, E., and Gu, S.G. (2014). Complex coding of endogenous siRNA, transcriptional silencing and H3K9 methylation on native targets of germline nuclear RNAi in C. elegans. BMC Genomics *15*, 1157.

Nishida, K.M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., Siomi, H., and Siomi, M.C. (2007). Gene silencing mechanisms mediated by Aubergine–piRNA complexes in *Drosophila* male gonad. RNA *13*, 1911–1922.

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J., et al. (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature *491*, 284–287.

Orgel, L.E., and Crick, F.H.C. (1980). Selfish DNA: the ultimate parasite. Nature *284*, 604–607.

Özata, D.M., Yu, T., Mou, H., Gainetdinov, I., Colpan, C., Cecchini, K., Kaymaz, Y., Wu, P.-H., Fan, K., Kucukural, A., et al. (2020). Evolutionarily conserved

pachytene piRNA loci are highly divergent among modern humans. Nat. Ecol. Evol. *4*, 156–168.

Pak, J., and Fire, A. (2007). Distinct Populations of Primary and Secondary Effectors During RNAi in C. elegans. Science *315*, 241–244.

Pane, A., Wehr, K., and Schüpbach, T. (2007). *zucchini* and *squash* encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. Dev. Cell *12*, 851–862.

Pane, A., Jiang, P., Zhao, D.Y., Singh, M., and Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. EMBO Journal *30*, 4601–4615.

Parhad, S.S., Tu, S., Weng, Z., and Theurkauf, W.E. (2017). Adaptive evolution leads to cross-species incompatibility in the piRNA transposon silencing machinery. Dev. Cell *43*, 60-70.

Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. Embo Journal *23*, 4727–4737.

Phillips, C.M., Brown, K.C., Montgomery, B.E., Ruvkun, G., and Montgomery, T.A. (2015). piRNAs and piRNA-Dependent siRNAs protect conserved and essential *C. elegans* genes from misrouting into the RNAi pathway. Dev. Cell *34*, 457–465.

Quénerch'du, E., Anand, A., and Kai, T. (2016). The piRNA pathway is developmentally regulated during spermatogenesis in *Drosophila*. RNA *22*, 1044–1054.

Reuter, M., Chuma, S., Tanaka, T., Franz, T., Stark, A., and Pillai, R.S. (2009). Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat. Struct. Mol. Biol. *16*, 639–646.

Reuter, M., Berninger, P., Chuma, S., Shah, H., Hosokawa, M., Funaya, C., Antony, C., Sachidanandam, R., and Pillai, R.S. (2011). Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature *480*, 264–267.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3′ UTR-directed primary piRNAs. Curr. Biol. *19*, 2066–2076.

Roovers, E.F., Rosenkranz, D., Mahdipour, M., Han, C.-T., He, N., Chuva de Sousa Lopes, S.M., van der Westerlaken, L.A.J., Zischler, H., Butter, F., Roelen, B.A.J., et al. (2015). Piwi proteins and piRNAs in mammalian oocytes and early embryos. Cell Reports *10*, 2069–2082.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell *127*, 1193–1207.

Sarkies, P., Selkirk, M.E., Jones, J.T., Blok, V., Boothby, T., Goldstein, B., Hanelt, B., Ardila-Garcia, A., Fast, N.M., Schiffer, P.M., et al. (2015). Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. PLoS Biol *13*, e1002061.

Schmidt, A., Palumbo, G., Bozzetti, M.P., Tritto, P., Pimpinelli, S., and Schäfer, U. (1999). Genetic and molecular characterization of *sting*, a gene involved in crystal formation and meiotic drive in the male germ line of *Drosophila melanogaster*. Genetics *151*, 749–760.

Schwarz, D.S., Tomari, Y., and Zamore, P.D. (2004). The RNA-Induced Silencing Complex Is a Mg2+-Dependent Endonuclease. Curr Biol 14, 787–791.

Senti, K.-A., Jurczak, D., Sachidanandam, R., and Brennecke, J. (2015). piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. Genes Dev *29*, 1747–1762.

Shen, E.-Z., Chen, H., Ozturk, A.R., Tu, S., Shirayama, M., Tang, W., Ding, Y.-H., Dai, S.-Y., Weng, Z., and Mello, C.C. (2018). Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the *C. elegans* germline. Cell *172*, 937-951.

Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. Cell *150*, 65–77.

Shoji, M., Tanaka, T., Hosokawa, M., Reuter, M., Stark, A., Kato, Y., Kondoh, G., Okawa, K., Chujo, T., Suzuki, T., et al. (2009). The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. Dev. Cell *17*, 775–787.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H.A. (2007). Secondary siRNAs Result from Unprimed RNA Synthesis and Form a Distinct Class. Science *315*, 244–247.

Skene, P.J., and Henikoff, S. (2015). A simple method for generating high-resolution maps of genome-wide protein binding. eLife *4*, e09225.

Stein, P., Rozhkov, N.V., Li, F., Cárdenas, F.L., Davydenko, O., Davydenk, O., Vandivier, L.E., Gregory, B.D., Hannon, G.J., and Schultz, R.M. (2015). Essential role for endogenous siRNAs during meiosis in mouse oocytes. PLoS Genet. *11*, e1005013.

Tan, M., Tol, H.T.A. van, Rosenkranz, D., Roovers, E.F., Damen, M.J., Stout, T.A.E., Wu, W., and Roelen, B.A.J. (2020). PIWIL3 forms a complex with TDRKH in mammalian oocytes. Cells *9*, 1356.

Tang, W., Seth, M., Tu, S., Shen, E.-Z., Li, Q., Shirayama, M., Weng, Z., and Mello, C.C. (2018). A sex chromosome piRNA promotes robust dosage compensation and sex determination in *C. elegans*. Dev. Cell *44*, 762-770.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320–324.

Vasiliauskaitė, L., Vitsios, D., Berrens, R.V., Carrieri, C., Reik, W., Enright, A.J., and O'Carroll, D. (2017). A MILI-independent piRNA biogenesis pathway empowers partial germline reprogramming. Nat Struct Mol Biol *24*, 604–606.

Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat. Struct. Mol. Biol. *19*, 773–781.

Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr. Biol. *18*, 861–867.

Wang, J., Saxe, J.P., Tanaka, T., Chuma, S., and Lin, H. (2009). Mili interacts with Tudor domain-containing protein 1 in regulating spermatogenesis. Curr. Biol. *19*, 640–644.

Wang, W., Yoshikawa, M., Han, B.W., Izumi, N., Tomari, Y., Weng, Z., and Zamore, P.D. (2014). The initial uridine of primary piRNAs does not create the tenth adenine that is the hallmark of secondary piRNAs. Mol. Cell *56*, 708–716.

Wang, W., Han, B.W., Tipping, C., Ge, D.T., Zhang, Z., Weng, Z., and Zamore, P.D. (2015). Slicing and binding by Ago3 or Aub trigger Piwi-bound piRNA production by distinct mechanisms. Mol. Cell *59*, 819–830.

Watanabe, T., Chuma, S., Yamamoto, Y., Kuramochi-Miyagawa, S., Totoki, Y., Toyoda, A., Hoki, Y., Fujiyama, A., Shibata, T., Sado, T., et al. (2010). MITOPLD Is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. Dev. Cell *20*, 364–375.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R.A., Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E.A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in *C. elegans*. Genes Dev. *28*, 783–796.

Weng, C., Kosalka, J., Berkyurek, A.C., Stempor, P., Feng, X., Mao, H., Zeng, C., Li, W.-J., Yan, Y.-H., Dong, M.-Q., et al. (2019). The USTC co-opts an ancient machinery to drive piRNA transcription in *C. elegans*. Genes Dev. *33*, 90–102.

Williams, Z., Morozov, P., Mihailovic, A., Lin, C., Puvvula, P.K., Juranek, S., Rosenwaks, Z., and Tuschl, T. (2015). Discovery and characterization of piRNAs in the human fetal ovary. Cell Reports *13*, 854–863.

Wit, E. de, Linsen, S.E.V., Cuppen, E., and Berezikov, E. (2009). Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Research *19*, 2064–2074.

Wu, P.-H., Fu, Y., Cecchini, K., Özata, D.M., Arif, A., Yu, T., Colpan, C., Gainetdinov, I., Weng, Z., and Zamore, P.D. (2020). The evolutionarily conserved piRNA-producing locus pi6 is required for male mouse fertility. Nat Genet *52*, 728–739.

Yu, B., Lin, Y.A., Parhad, S.S., Jin, Z., Ma, J., Theurkauf, W.E., Zhang, Z.Z., and Huang, Y. (2018). Structural insights into Rhino-Deadlock complex for germline piRNA cluster specification. EMBO Rep. *19*, e45418.

Yuan, Y.-R., Pei, Y., Ma, J.-B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H.-Y., Dauter, Z., Tuschl, T., and Patel, D.J. (2005). Crystal Structure of A. aeolicus Argonaute, a Site-Specific DNA-Guided Endoribonuclease, Provides Insights into RISC-Mediated mRNA Cleavage. Mol Cell *19*, 405–419.

Yu, T., Koppetsch, B.S., Pagliarani, S., Johnston, S., Silverstein, N.J., Luban, J., Chappell, K., Weng, Z., and Theurkauf, W.E. (2019). The piRNA response to retroviral invasion of the koala genome. Cell *179*, 632-643.

Zhang, D., Tu, S., Stubna, M., Wu, W.-S., Huang, W.-C., Weng, Z., and Lee, H.-C. (2018). The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science *359*, 587–592.

Zhang, Z., Xu, J., Koppetsch, B.S., Wang, J., Tipping, C., Ma, S., Weng, Z., Theurkauf, W.E., and Zamore, P.D. (2011). Heterotypic piRNA ping-pong requires Qin, a protein with both E3 Ligase and Tudor domains. Mol. Cell *44*, 572–584.

Zhao, S., Gou, L.-T., Zhang, M., Zu, L.-D., Hua, M.-M., Hua, Y., Shi, H.-J., Li, Y., Li, J., Li, D., et al. (2013). piRNA-Triggered MIWI Ubiquitination and Removal by APC/C in Late Spermatogenesis. Dev Cell *24*, 13–25.

Chapter 2

SNPC-1.3 is a sex-specific transcription factor that drives male piRNA expression in *C. elegans*

2.1 Citation

Charlotte P. Choi*[1], Rebecca J. Tay*[1], Margaret R. Starostik[1], Suhua Feng[2,3], James J. Moresco[4], Brooke E. Montgomery[5], Emily Xu[1], Maya A. Hammonds[1], Michael C. Schatz[1,6], Taiowa A. Montgomery[5], John R. Yates III[7], Steven E. Jacobsen[2,8], John K. Kim[1,9] (2021). SNPC-1.3 is a sex-specific transcription factor that drives male piRNA expression in *C. elegans*. eLife *10*, e60681.

2.2 Author Contributions

Charlotte P Choi: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing

Rebecca J Tay: Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing

John K Kim: Conceptualization, Supervision, Funding acquisition, Methodology, Project administration, Writing - review and editing

Margaret R. Starostik: Data curation, Formal analysis, Visualization, Methodology, Writing - original draft

Michael C Schatz: Formal analysis, Supervision, Methodology

Taiowa A. Montgomery: Formal analysis, Methodology

Suhua Feng, James J. Moresco, Brooke E. Montgomery, Emily Xu, Maya A. Hammonds, Steven E. Jacobsen: Methodology

2.3 Abstract

Piwi-interacting RNAs (piRNAs) play essential roles in silencing repetitive elements to promote fertility in metazoans. Studies in worms, flies, and mammals reveal that piRNAs are expressed in a sex-specific manner. However, the mechanisms underlying this sex-specific regulation are unknown. Here we identify SNPC-1.3, a male germline-enriched variant of a conserved subunit of the small nuclear RNA activating protein complex, as a male-specific piRNA transcription factor in *C. elegans*. SNPC-1.3 colocalizes with the core piRNA transcription factor, SNPC-4, in nuclear foci of the male germline. Binding of SNPC-1.3 at male piRNA loci drives spermatogenic piRNA transcription and requires SNPC-4. Loss of *snpc-1.3* leads to depletion of male piRNAs and defects in male-dependent fertility. Furthermore, TRA-1, a master regulator of sex determination, binds to the *snpc-1.3* promoter and represses its expression during oogenesis. Loss of TRA-1 targeting causes ectopic expression of *snpc-1.3* and male piRNAs during oogenesis. Thus, sexually dimorphic regulation of *snpc-1.3* expression coordinates male and female piRNA expression during germline development.

2.4 Introduction

Piwi-interacting RNAs (piRNAs), a distinct class of small noncoding RNAs, function to preserve germline integrity (Batista et al., 2008; Carmell et al., 2007; Cox et al., 1998; Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2008; Lin and Spradling, 1997; Wang and Reinke, 2008). In *Drosophila*, mutation of any of the three Piwi genes (*piwi*, *aub*, *ago3*) results in rampant activation of transposons in the germline and severe defects in fertility (Brennecke et al., 2007; Harris and Macdonald, 2001; Lin and Spradling, 1997; Vagin et al., 2006). In *M. musculus*, mutation of the Piwi protein MIWI leads to the misregulation of genes involved in germ cell development, defective gametogenesis, and sterility (Deng and Lin, 2002; Zhang et al., 2015b). *C. elegans* piRNAs can be inherited across multiple generations and trigger the transgenerational silencing of foreign elements such as transgenes. Disruption of this inheritance results in eventual germline collapse and sterility, known as the germline mortal phenotype (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). Taken together, piRNAs are essential to preserve germline integrity and ensure the reproductive capacity in metazoans.

Loss of the piRNA pathway can have distinct consequences between the sexes and across developmental stages. Many species show sex-specific expression of piRNAs (Armisen et al., 2009; Billi et al., 2013; Williams et al., 2015; Yang et al., 2013; Zhou et al., 2010). Demonstrated by hybrid dysgenesis, the identity of female, but not male, piRNAs in flies is important for fertility (Brennecke et al., 2008). In contrast, the piRNA pathway in mammals appears to be dispensable for female fertility (Carmell et al., 2007; Murchison et al., 2007), but

distinct subsets of piRNAs are required for specific stages of spermatogenesis (Aravin et al., 2003; Aravin et al., 2006; Carmell et al., 2007; Di Giacomo et al., 2013; Gainetdinov et al., 2018; Girard et al., 2006; Grivna et al., 2006; Kuramochi-Miyagawa et al., 2008; Li et al., 2013). In worms, most piRNAs are uniquely enriched in either the male or female germline (Billi et al., 2013; Kato et al., 2009). Nevertheless, in all of these contexts, how the specific expression of different piRNA subclasses is achieved is poorly understood.

piRNA biogenesis is strikingly diverse across organisms and tissue types. In the *Drosophila* germline, piRNA clusters are found within pericentromeric or telomeric heterochromatin enriched for H3K9me3 histone modifications. The HP1 homolog Rhino binds to H3K9me3 within most of these piRNA clusters and recruits Moonshiner, a paralog of the basal transcription factor TFIIA, which, in turn, recruits RNA polymerase II (Pol II) to enable transcription within heterochromatin (Andersen et al., 2017; Chen et al., 2016; Klattenhoff et al., 2009; Mohn 2014; et al., Pane et al., 2011). Two waves of piRNA expression occur in mouse testes: pre-pachytene piRNAs are expressed in early spermatogenesis and silence transposons, whereas pachytene piRNAs are expressed in the later stages of meiosis and have unknown functions. While the mechanisms of pre-pachytene piRNA transcription remain elusive, pachytene piRNAs require the transcription factor A-MYB, along with RNA Pol II (Li et al., 2013).

In *C. elegans*, SNPC-4 is essential for the expression of piRNAs in the germline (Kasper et al., 2014). SNPC-4 is the single *C. elegans* ortholog of mammalian SNAPC4, the largest DNA binding subunit of the small nuclear RNA

(snRNA) activating protein complex (SNAPc). A complex of SNAPC4, SNAPC1, and SNAPC3 binds to the proximal sequence element (PSE) of snRNA loci to promote their transcription (Henry et al., 1995; Jawdekar and Henry, 2008; Ma and Hernandez, 2002; Su et al., 1997; Wong et al., 1998; Yoon et al., 1995). SNPC-4 occupies transcription start sites of other classes of noncoding RNAs across various *C. elegans* tissue types and developmental stages (Kasper et al., 2014; Weng et al., 2019). Furthermore, piRNA biogenesis factors PRDE-1, TOFU-4, and TOFU-5 are expressed in germ cell nuclei and interact with SNPC-4 at clusters of piRNA loci (Goh et al., 2014; Kasper et al., 2014; Weick et al., 2014; Weng et al., 2019). These data suggest that SNPC-4 has been co-opted by germline-specific factors to transcribe piRNAs.

The vast majority of the ~15,000 piRNAs in *C. elegans* are encoded within two large megabase genomic clusters on chromosome IV (Das et al., 2008; Ruby et al., 2006). Each piRNA locus encodes a discrete transcriptional unit that is individually transcribed as a short precursor by Pol II (Gu et al., 2012; Cecere et al., 2012; Billi et al., 2013). Processing of precursors yields mature piRNAs that are typically 21 nucleotides (nt) in length and strongly enriched for a 5' uracil (referred to as 21U-RNAs). Transcription of these piRNAs requires a conserved 8 nt core motif (NNGTTTCA) within their promoters (Billi et al., 2013; Cecere et al., 2012; Ruby et al., 2006). piRNAs enriched during spermatogenesis are associated with a cytosine at the 5' most position of the core motif (CNGTTTCA); mutation of cytosine to adenine at this position results in ectopic expression of normally male-enriched piRNAs during oogenesis. In contrast, genomic loci expressing piRNAs

enriched in the female germline show no discernable nucleotide bias at the 5' position (Billi et al., 2013). While differences in *cis*-regulatory sequences contribute to the sexually dimorphic nature of piRNA expression, sex-specific piRNA transcription factors that drive distinct subsets of piRNAs in the male and female germlines remain to be identified.

Here, we demonstrate that SNPC-1.3, an ortholog of human SNAPC1, is required specifically for male piRNA expression. Furthermore, TRA-1, a master regulator of sex-determination, transcriptionally represses *snpc-1.3* during oogenesis to restrict its expression to the male germline. Taken together, our study reports the first example of a sex-specific piRNA transcription factor that drives the expression of male-specific piRNAs.

2.5 Results

2.5.1 SNPC-4 is a component of the core piRNA transcription complex that drives all piRNA expression.

SNPC-4-specific foci are present in both male and female germ cell nuclei (Kasper et al., 2014), but the role of SNPC-4 in the male germline is not well understood. I hypothesized that SNPC-4 is required for piRNA biogenesis in both the male and female germlines. To test this, we conditionally depleted the SNPC-4 protein using the auxin-inducible degradation system (Zhang et al., 2015a) (Figure 2.4). We added an auxin-inducible degron (AID) to the C-terminus of SNPC-4 using CRISPR/Cas9 genome engineering, and crossed this strain into worms expressing TIR1 under the germline promoter, *sun-1*. TIR1 is a plant-specific F-box protein that mediates the rapid degradation of *C. elegans* proteins

tagged with an AID in the presence of the phytohormone auxin. Thus, addition of auxin to the *snpc-4::aid; Psun-1::TIR1* strain is expected to degrade SNPC-4::AID, whereas strains with *snpc-4::aid* alone serve as a negative control; under these conditions, we examined a panel of spermatogenesis- and oogenesis-enriched piRNAs (Billi et al., 2013) during spermatogenesis and oogenesis. Unless otherwise stated, spermatogenesis and oogenesis stages will correspond to time points taken at 48 h and 72 h, respectively, post L1 hatching at 20°C. Worms depleted of SNPC-4 showed decreased expression of both spermatogenesis- and oogenesis-enriched piRNAs during spermatogenesis and oogenesis time points, respectively (Figure 2.1A), confirming that SNPC-4 is a core piRNA transcription factor required for all piRNA expression.

Given that SNPC-4 activates transcription of piRNAs in both sexes, I hypothesized that sex-specific cofactors might associate with SNPC-4 to regulate sexually dimorphic piRNA expression. To test this hypothesis, I leveraged genetic backgrounds that masculinize or feminize the germline. Specifically, I used *him-8(-)* mutants, which have a higher incidence of males (~30% males compared to <0.5% spontaneous males in the wild-type hermaphrodite population) (Hodgkin et al., 1979), and *fem-1(-)* mutants, which are completely feminized when grown at 25°C (Doniach and Hodgkin, 1984). I introduced a C-terminal 3xFlag tag sequence at the endogenous *snpc-4* locus using CRISPR/Cas9 genome editing (Paix et al., 2015) and performed immunoprecipitation of SNPC-4::3xFlag followed by mass spectrometry. PRDE-1 and TOFU-5 co-purified with SNPC-4::3xFlag in both *him-8(-)* and *fem-1(-)* mutants, suggesting that these known piRNA biogenesis factors

exist as a complex in both male and female germlines (Figure 2.1B, 2.2; and Table 2.1). While a single worm ortholog, SNPC-4, exists for human SNAPC4, the *C. elegans* genome encodes 4 homologs of human SNAPC1 (worm SNPC-1.1, -1.2, -1.3, and -1.5) and 4 homologs of human SNAPC3 (worm SNPC-3.1, -3.2, -3.3, and -3.4) (Figure 2.1B) (Li et al., 2004). From our mass spectrometry analysis, 6 out of the 8 *C. elegans* homologs of SNAPC1 and SNAPC3 co-purified with SNPC-4::3xFlag from both *him-8(-)* and *fem-1(-)* genetic backgrounds (Figure 2.1B, 2.2). These results revealed that SNPC-4 interacts with both snRNA and piRNA transcriptional machinery.

2.5.2 SNPC-1.3 interacts with the core piRNA biogenesis factor SNPC-4 during spermatogenesis.

I also identified proteins that co-purified with SNPC-4::3xFlag from *him-8(-)*, but not *fem-1(-)* mutants. I was particularly interested in SNPC-1.3 because of its homology to the mammalian SNAPC1 subunit of the snRNA transcription complex. We confirmed that SNPC-1.3 interacts with SNPC-4 by using CRISPR/Cas9 genome editing to generate an endogenously tagged *snpc-1.3::ollas* strain. We then crossed *snpc-1.3::ollas* into the *snpc-4::3xflag* strain and performed immunoprecipitation with anti-Flag antibodies. In agreement with the mass spectrometry data, SNPC-4::3xFlag and SNPC-1.3:Ollas interacted robustly during spermatogenesis. The interaction was detectable at a much lower level during oogenesis (Figure 2.3). The reciprocal co-immunoprecipitation of SNPC-1.3::3xFlag followed by western blotting for SNPC-4::Ollas confirmed this

biochemical interaction (Figure 2.6), suggesting that SNPC-1.3 forms a complex with the previously characterized piRNA biogenesis factor SNPC-4.

2.5.3 SNPC-1.3 is enriched in the male germline.

To determine whether SNPC-1.3 expression is restricted to the germline, we first examined *snpc-1.3* mRNA levels during early spermatogenesis (36 h post L1 hatching) in worms fed *glp-1* RNAi, which abrogates germline development. The mRNA of *snpc-4,* which is highly expressed in the germline (Kasper et al., 2014), was used as a control. Knockdown of *glp-1* mRNA markedly reduced both *snpc-1.3* and *snpc-4* mRNAs. SNPC-1.3::3xFlag protein expression was also reduced in a *glp-4* temperature-sensitive mutant, which fails to develop fully-expanded germlines at 25°C (Beanan and Strome, 1992), suggesting that SNPC-1.3 is predominantly expressed in the germline (Figure 2.7A).

To examine differential *snpc-1.3* expression between the sexes, we measured *snpc-1.3* mRNA levels in *him-8(-)* males and *fem-1(-)* females. The expression of *snpc-1.3* mRNA was greatly enriched in *him-8(-)* relative to *fem-1(-)*, while *snpc-4* mRNA did not show any differential expression. At the protein level, SNPC-1.3::3xFlag was also highly enriched in males as compared to females by western blotting (Figure 2.7B).

SNPC-4, along with other piRNA factors, such as PRDE-1, localize to one or two foci in each germline nuclei (Kasper et al., 2014, Weick et al., 2014, Weng et al., 2019). Given that SNPC-1.3 is present in a complex with SNPC-4 (Figure 2.2, 2.3), we hypothesized that SNPC-1.3 might show a similar localization pattern to these other piRNA factors. To examine the subcellular localization of SNPC-1.3,

we performed immunofluorescence in *snpc-4::3xflag; snpc-1.3::ollas* adult males and hermaphrodites. In the male germline, SNPC-1.3::Ollas colocalized with SNPC-4::3xFlag in the same nuclear foci (Figure 2.7). In contrast, no SNPC-1.3::Ollas signal was detected above background in hermaphrodites (Figure 2.7). Taken together, these data indicate that SNPC-1.3 co-localizes with SNPC-4 specifically in the male germline.

2.5.4 SNPC-1.3 is required for transcription of male piRNAs.

Given the prominent interaction between SNPC-1.3 and SNPC-4 in the male germline (Figure 2.3), I hypothesized that SNPC-1.3 might be required for piRNA expression during spermatogenesis. To test this hypothesis, I generated a *snpc-1.3* null allele by introducing mutations that result in a premature stop codon located 8 amino acids away from the start codon at the *snpc-1.3* locus. I examined spermatogenesis in hermaphrodites and *him-8(-)* males, and examined oogenesis in adult hermaphrodites and *fem-1(-)* females. As a control, I analyzed the loss-of-function mutant of the *C. elegans* Piwi protein, *prg-1(-)*, which almost completely lacked male and female piRNAs (Figure 2.8A), as expected. Levels of male piRNAs were dramatically reduced in *snpc-1.3(-)* hermaphrodites during spermatogenesis and in *him-8(-); snpc-1.3(-)* males, whereas female piRNAs were largely unaltered in *snpc-1.3(-)* adult hermaphrodites and in *fem-1(-); snpc-1.3(-)* females (Figure 2.8A, B). Unexpectedly, female piRNAs were also moderately upregulated by at least 2-fold in *snpc-1.3(-)* mutants undergoing spermatogenesis and in *him-8(-); snpc-1.3(-)* males. These findings suggest that, in addition to activating male piRNAs, SNPC-1.3 suppresses the expression of female piRNAs

in the male germline, possibly by preferentially recruiting core factors such as SNPC-4 to male piRNA loci. As SNPC-4 is known to activate transcription of snRNAs as well as piRNAs (Kasper et al., 2014), I asked whether SNPC-1.3 is also required for transcribing snRNAs. To test this, we measured U1 snRNA levels in hermaphrodite adults after RNAi-mediated knockdown of *snpc-1.3*. In contrast to the reduction of U1 observed in *snpc-4* RNAi, U1 levels were not significantly altered when *snpc-1.3* was depleted (Figure 2.9), suggesting that, unlike SNPC-4, SNPC-1.3 is likely specific to the transcription of male piRNAs and does not play a role in snRNA transcription.

To extend these findings, we identified piRNAs enriched during spermatogenesis and oogenesis by small RNA-seq in wild-type worms. Using a 1.2-fold threshold and false discovery rate (FDR) of ≤ 0.05, a total of 6,368 out of 14,714 piRNAs on chromosome IV were differentially expressed (Figure 2.10A). Among these, 4,060 piRNAs were upregulated during spermatogenesis (hereafter referred to as male piRNAs) and 2,308 piRNAs were upregulated during oogenesis, which we define as female piRNAs. We compared this dataset with our previous study that identified and categorized spermatogenesis- and oogenesis-enriched piRNAs, as well as piRNAs that were not statistically enriched (NE) either during oogenesis or spermatogenesis (Billi et al., 2013). Most male piRNAs identified in this study were also identified in our previous study (82%; 3,316/4,060) (Figure 2.10A). Next, we investigated how loss of *snpc-1.3* affects global piRNA expression by performing small RNA-seq in wildtype versus *snpc-1.3(-)* mutants during spermatogenesis. We identified 3,601 piRNAs that were downregulated in

a *snpc-1.3(-)* mutant compared to wildtype (Figure 2.10B). Of these, 3,002 overlapped with spermatogenesis-enriched piRNAs identified in our previous study (Billi et al., 2013) (Figure 2.10B). 85% (3,452/4,060) of male piRNAs were also depleted in *snpc-1.3(-)* mutants, suggesting that male piRNAs are regulated by SNPC-1.3 (Figure 2.11A). Consistent with our Taqman analysis (Figure 2.8A–B), 73% (1,687/2,308) of oogenesis-enriched piRNAs identified in our study were significantly upregulated in *snpc-1.3(-)* mutants during spermatogenesis (Figure 2.11B).

We next analyzed the genomic loci of male piRNAs and *snpc-1.3*-dependent piRNAs. As expected, the intersection of these two piRNA subsets displayed strong enrichment for the 8 nt core motif and the 5'-most position of this core motif was enriched for cytosine (CNGTTTCA) (Figure 2.11A). In contrast, the core motif found upstream of female piRNAs upregulated upon loss of *snpc-1.3* displayed a much weaker bias for the 5' cytosine (Figure 2.11B). These observations validate our previous findings that male and female core motifs are distinct (Billi et al., 2013). Taken together, these data indicate that SNPC-1.3 is required for male piRNA expression.

2.5.5 SNPC-1.3 binds male piRNA loci in a SNPC-4-dependent manner.

Given that SNPC-1.3 interacts with SNPC-4 and is required for expression of male piRNAs, we hypothesized that SNPC-1.3 might bind male piRNA loci in association with SNPC-4. To test this, we performed ChIP-qPCR to investigate SNPC-1.3 occupancy at regions of high piRNA density within the two large piRNA clusters on chromosome IV; an intergenic region lacking piRNAs served as a

control. To determine whether SNPC-1.3 binding was dependent on SNPC-4, we again used the auxin-inducible degradation system to deplete SNPC-4 in the *snpc-1.3::3xflag* strain for 4 hours prior to our spermatogenesis time point. In the presence of SNPC-4 expression, SNPC-1.3 was enriched at both piRNA clusters, albeit to a lesser degree at the small cluster, and this enrichment was lost upon SNPC-4 depletion (Figure 2.12A). These data indicate that SNPC-1.3 binds piRNA loci during spermatogenesis in a SNPC-4-dependent manner *in vivo*.

To examine the genome-wide binding profile of SNPC-1.3 and its dependency on SNPC-4, we performed ChIP-seq of N2, *snpc-1.3::3xflag*, and *snpc-1.3::3xflag; snpc-4::aid; Psun-1::TIR1* worms during spermatogenesis (Figure 2.12). Consistent with our ChIP-qPCR results, SNPC-1.3 binds piRNA clusters in a SNPC-4-dependent manner (Figure 2.12B). By quantifying the SNPC-1.3 signal over consecutive, non-overlapping 1 kb bins across the entire genome, we identified 691 1 kb regions within the chromosome IV piRNA clusters that were enriched for SNPC-1.3 in *snpc-1.3::3xflag* compared to N2 (Figure 2.12C). Relative to *snpc-1.3::3xflag*, worms depleted of SNPC-4 showed loss of SNPC-1.3 in 749 1 kb regions on chromosome IV piRNA clusters (Figure 2.12D). Furthermore, SNPC-1.3 enrichment ($p< 2.2 \times 10^{-16}$) and depletion ($p< 2.2 \times 10^{-16}$) were specific to the piRNA clusters on chromosome IV, and more than half (393/691) of the SNPC-1.3-enriched regions in *snpc-1.3::3xflag* worms were depleted upon degradation of SNPC-4 (Figure 2.12C–D).

To determine whether SNPC-1.3 preferentially binds male piRNA loci, we characterized the SNPC-1.3 signal around individual 5’ nucleotides of mature

piRNAs. Again, we classified piRNAs as male, female, or not significantly enriched (NE) in either sex, based on our small RNA-seq analysis in wild-type hermaphrodites during spermatogenesis and oogenesis (Figure 2.10A). SNPC-1.3 binding at male piRNA loci was most enriched just upstream of the piRNA 5' nucleotide, which overlaps the conserved core motif (Figure 2.13A). This binding profile was very distinct for 1 kb bins that contained only male piRNAs (Figure 2.13B). Upon depletion of SNPC-4, this peak in male piRNAs was lost (Figure 2.13A). Although the binding profiles for individual female piRNAs exhibited more variability, there was little evidence for SNPC-1.3 binding and dependency on SNPC-4 at female loci (Figure 2.13A). Compared to the binding profile in male piRNA loci, SNPC-1.3 binding was observed to a lesser extent in non-enriched piRNAs (Figure 2.13A). Taken together, these observations indicate that SNPC-1.3 requires the core piRNA factor SNPC-4 to bind the piRNA clusters during spermatogenesis.

2.5.6 TRA-1 represses snpc-1.3 and male piRNA expression during oogenesis.

As male piRNA expression and SNPC-1.3 protein expression are largely restricted to the male germline, we asked how *snpc-1.3* mRNA expression is regulated across development. *C. elegans* hermaphrodites produce sperm during the L4 stage and transition to producing oocytes as adults. To understand the mRNA expression profile of *snpc-1.3* relative to *snpc-4* and other developmentally regulated genes, I performed qRT-PCR across hermaphrodite development. *snpc-4* mRNA is expressed at low levels during spermatogenesis, but dramatically

increases during oogenesis (Figure 2.6, 2.14). These data suggest that low levels of SNPC-4 are sufficient for activating male piRNA biogenesis during spermatogenesis. Consistent with SNPC-1.3 protein expression (Figure 2.3), *snpc-1.3* mRNA levels peak in L3 to early L4 stages, during spermatogenesis (Figure 2.14). Given that *snpc-1.3* expression across development is regulated at the mRNA level, we examined the sequences upstream of the *snpc-1.3* coding region to identify potential *cis*-regulatory motifs. Less than 200 bp upstream of the *snpc-1.3* start codon, I identified three consensus binding sites for TRA-1 (Figure 2.15), a transcription factor that controls the transition from spermatogenesis to oogenesis (Berkseth et al., 2013; Clarke and Berg, 1998; Zarkower and Hodgkin, 1993).

In the germline, TRA-1, a Gli family zinc-finger transcription factor, controls the sperm-to-oocyte decision by repressing both *fog-1* and *fog-3,* which are required for controlling sexual cell fate (Berkseth et al., 2013; Chen and Ellis, 2000; Lamont and Kimble, 2007; Zarkower and Hodgkin, 1993). Loss-of-function *tra-1* hermaphrodites exhibit masculinization of the female germline and develop phenotypically male-like traits (Hodgkin, 1987). I used RNAi to knock down *tra-1* and observed significant ectopic upregulation of *snpc-1.3* mRNA during oogenesis (Figure 2.16). However, this upregulation of *snpc-1.3* expression could be an indirect effect of masculinization of the germline. Therefore, to test whether TRA-1 directly regulates *snpc-1.3*, I generated strains harboring mutations at the three TRA-1 binding sites (*tbs*) in the endogenous *snpc-1.3* promoter. Specifically, I mutated one (*1xtbs*), two (*2xtbs*), or all three (*3xtbs*) consensus TRA-1 binding

motifs (Figure 2.15). Disruption of the TRA-1 binding sites led to reduced TRA-1::3xFlag binding upstream of *snpc-1.3* as revealed by ChIP-seq, with the *3xtbs* mutant showing the greatest reduction of binding (Figure 2.15). In addition, *snpc-1.3* mRNA levels were highly upregulated when multiple TRA-1 binding sites were mutagenized (Figure 2.16), consistent with TRA-1 directly repressing *snpc-1.3* transcription during oogenesis. To confirm that SNPC-1.3 protein expression was also elevated in TRA-1 binding site mutants, we used CRISPR/Cas9 engineering to add a C-terminal 3xFlag tag at the *snpc-1.3* locus in *snpc-1.3(2xtbs)* mutants. Indeed, SNPC-1.3::3xFlag showed increased expression in the *snpc-1.3::3xFlag(2xtbs)* mutant during spermatogenesis and especially oogenesis (Figure 2.16). Taken together, these findings demonstrate TRA-1 binds to the *snpc-1.3* promoter to repress its transcription during oogenesis.

Given that *snpc-1.3* is robustly de-repressed during oogenesis in TRA-1 binding site mutants, I hypothesized that male piRNAs would also be ectopically upregulated during oogenesis. To test this, we performed small RNA-seq and compared piRNA levels in wild-type and *snpc-1.3(2xtbs)* worms during oogenesis. Using a 1.2-fold threshold and FDR of ≤ 0.05, we observed 1,370 piRNAs in *snpc-1.3(2xtbs)* mutants that were upregulated compared to wildtype (Figure 2.17A). The majority of these upregulated piRNAs overlap with the male piRNAs that we identified in wild-type hermaphrodites (Figure 2.17A). I also confirmed this result by Taqman qPCR analysis, which showed that male piRNAs were significantly upregulated in *snpc-1.3(2xtbs)* and *snpc-1.3(3xtbs)* mutants compared to wildtype during oogenesis (Figure 2.17B). Taken together, these data suggest that TRA-1

directly binds to *tbs* sites in the *snpc-1.3* promoter to repress its transcription and consequently, male piRNA expression during oogenesis.

Our data showed that female piRNAs are inappropriately upregulated during spermatogenesis upon loss of *snpc-1.3* (Figure 2.8). Consistent with this result, female piRNAs show reduced expression during oogenesis upon upregulation of SNPC-1.3 expression in *snpc-1.3(2xtbs)* and *snpc-1.3(3xtbs)* mutants compared to wildtype (Figure 2.17B). I posit that SNPC-1.3 plays a direct role in activating male piRNA transcription, while indirectly limiting female piRNA transcription by sequestering core piRNA transcription factors to male piRNA loci.

2.5.7 SNPC-1.3 is critical for male fertility.

Given the global depletion of male piRNAs in *snpc-1.3(-)* mutants and the progressive fertility defects seen in *prg-1(-)* mutants (Batista et al., 2008; Wang and Reinke, 2008), we hypothesized that *snpc-1.3(-)* worms might also show fertility defects. Indeed, *snpc-1.3(-)* hermaphrodites exhibited significantly reduced fertility compared to wildtype when grown at 25°C (Figure 2.18A). To address whether this decreased fertility was due to defects during spermatogenesis or oogenesis, we compared brood sizes from crosses of *fem-1(-)* females and *him-8(-)* males with or without *snpc-1.3*. Compared to *him-8(-)* males, *him-8(-); snpc-1.3(-)* males generated significantly smaller brood sizes when crossed with *fem-1(-)* females; in contrast, *fem-1(-); snpc-1.3(-)* and *fem-1(-)* females generated similar brood sizes when crossed with *him-8(-)* males (Figure 2.18B). As an orthogonal test, we crossed hermaphrodites to transgenic males expressing a fluorescent marker to facilitate counting of cross progeny. These transgenic males

encode a reporter gene, *Pcol-19::gfp*, which drives GFP expression in the cuticle (Figure 2.19A). All *Pcol-19::gfp; snpc-1.3(-)* males produced fewer GFP+ progeny than wild-type *Pcol-19::gfp* males, whereas wild-type or *snpc-1.3(-)* hermaphrodites generated similar numbers of GFP+ progeny when crossed with wild-type *Pcol-19::gfp* males (Figure 2.19A). These results suggest that the reduced fertility of *snpc-1.3(-)* mutants likely reflect defects during spermatogenesis.

To investigate the cause of *snpc-1.3*-dependent loss of male fertility, I examined spermiogenesis and sperm morphology in *snpc-1.3(-)* males. After meiotic differentiation in the male germline, male spermatids are induced by ejaculation and undergo spermiogenesis, a process that converts immature spermatids to motile sperm with a functioning pseudopod. Spermiogenesis can be induced *in vitro* by isolating spermatids directly from males and treating them with pronase (Shakes and Ward, 1989). Males lacking *prg-1* still generate differentiated spermatids, but rarely produce normal pseudopodia upon activation (Figures 2.20) (Wang and Reinke, 2008). Similar to *prg-1(-)* mutants, *snpc-1.3(-)* spermatids were rarely able to form normal pseudopodia. In contrast, *snpc-1.3(3xtbs)* sperm formed normal pseudopodia at a frequency similar to wildtype (Figures 2.20). In addition, many of the *snpc-1.3(-)* spermatids resembled sperm undergoing intermediate stages of spermiogenesis. Spermiogenesis, *in vivo*, starts off with spherical spermatids that enter into an intermediate stage characterized by the growth of spiky protrusions. This stage is then followed by fusion of the spiky protrusions into a motile pseudopod (Figure 2.21). To understand the dynamics of *snpc-1.3(-)*

sperm progression through spermiogenesis, I treated spermatids with pronase and observed each activated spermatid over time. Wild-type spermatids spent an average of 6.2 min ± 4.5 min in the intermediate state before polarization and pseudopod development. In contrast, *snpc-1.3(-)* spermatids occupied the intermediate state for a significantly shorter period of time (2.9 min ± 3.7 min, $p<0.05$; Student's t test) before forming pseudopods. By tracking each individual spermatid across spermiogenesis, I found most *snpc-1.3(-)* spermatids were unable to sustain pseudopod growth. While wild-type spermatids exhibited pseudopod growth and motility for an average of 24 min ± 10.35 min, *snpc-1.3(-)* spermatids sustained growth for a significantly shorter period of time (7.3 min ± 5.7 min, $p<0.05$; Student's t test) before becoming immotile (Figure 2.21). These results indicate that *spnc-1.3(-)* males have defective spermatogenesis processes and exhibit similar fertility defects as *prg-1(-)* mutants.

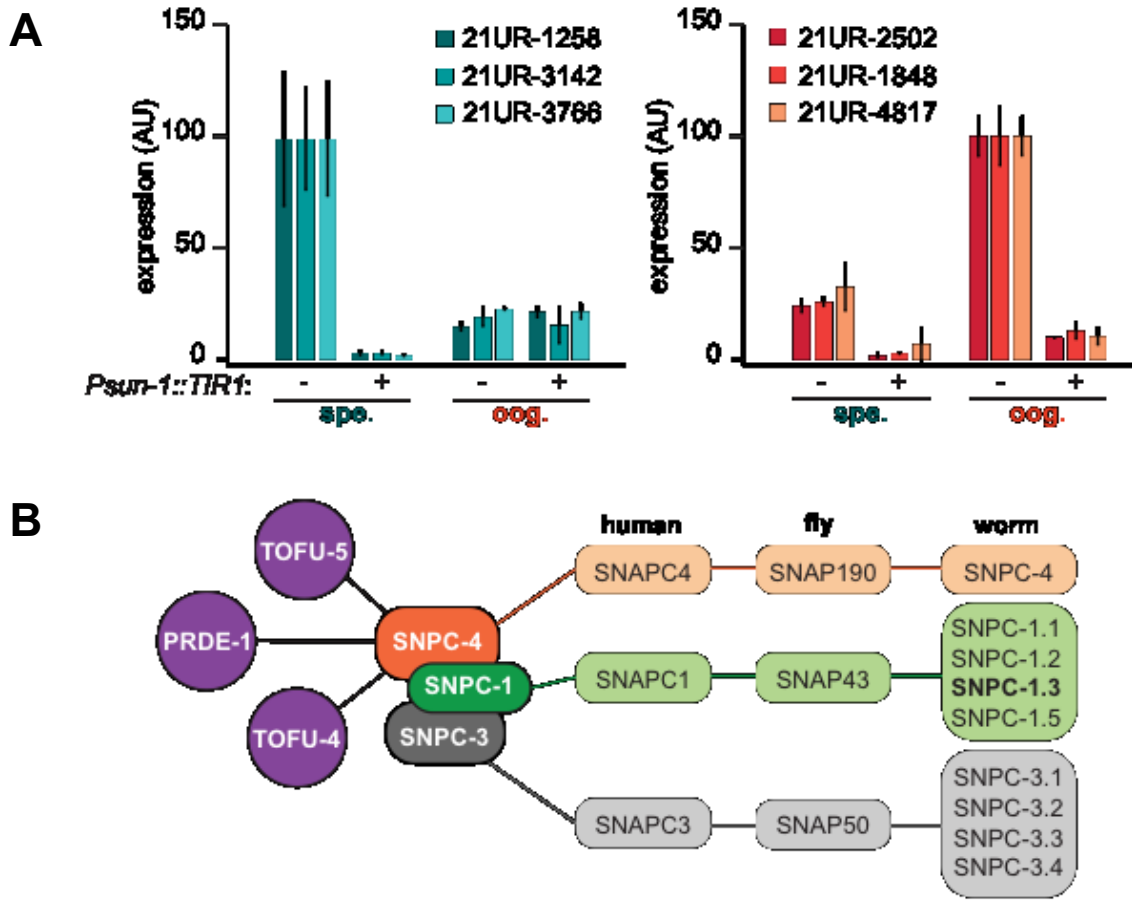

Figure 2.1 SNPC-4 and SNPC-1.3 are part of the male piRNA transcription complex.

(A) SNPC-4 is required for both male and female piRNA expression. Taqman qPCR of male (left) and female (right) piRNAs normalized to U18 small nucleolar RNA in *snpc-4::aid* (denoted as '-') and *snpc-4::aid; Psun-1::TIR1* (denoted as '+') worms. Both genotypes were placed on auxin, and collected during spermatogenesis (spe., 48 h) and oogenesis (oog., 72 h). Error bars: ± SD from two technical replicates. (B) Schematic highlights the conservation of SNAPc homologs from *C. elegans, D. melanogaster, and H. sapiens* and catalogs all SNPC-4 (orange) interacting partners from previous work (Weick et al., 2014; Weng et al., 2019) or from our own analysis. Known piRNA biogenesis factors (purple), SNPC-1 paralogs (green), and SNPC-3 paralogs (grey) are indicated.

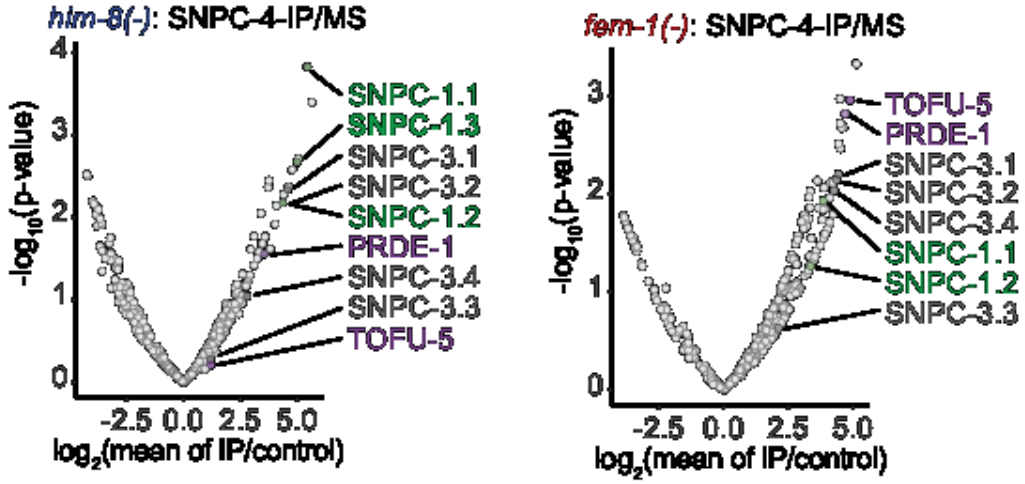

Figure 2.2. SNPC-1.3 interacts with SNPC-4 in only *him-8(-)* mutants.

Volcano plots showing enrichment values of IP of SNPC-4 over control (control: *him-8(-)* mutants for top panel or *fem-1(-)* mutants for bottom panel) and analogous significance values for proteins that co-purified with SNPC-4::3xFlag from (top) *him-8(-)* mutants or (bottom) *fem-1(-)* mutants (n = 2 biological replicates). piRNA biogenesis factors (purple), SNPC-1 paralogs (green), and SNPC-3 paralogs (dark grey) are labeled in Figure 1B. Although SNPC-3.1 and SNPC-3.2 are reported to have the same amino acid sequence, we have picked up differential peptide coverage in the *fem-1(-)* mutant for these two proteins and represented them as two different data points.

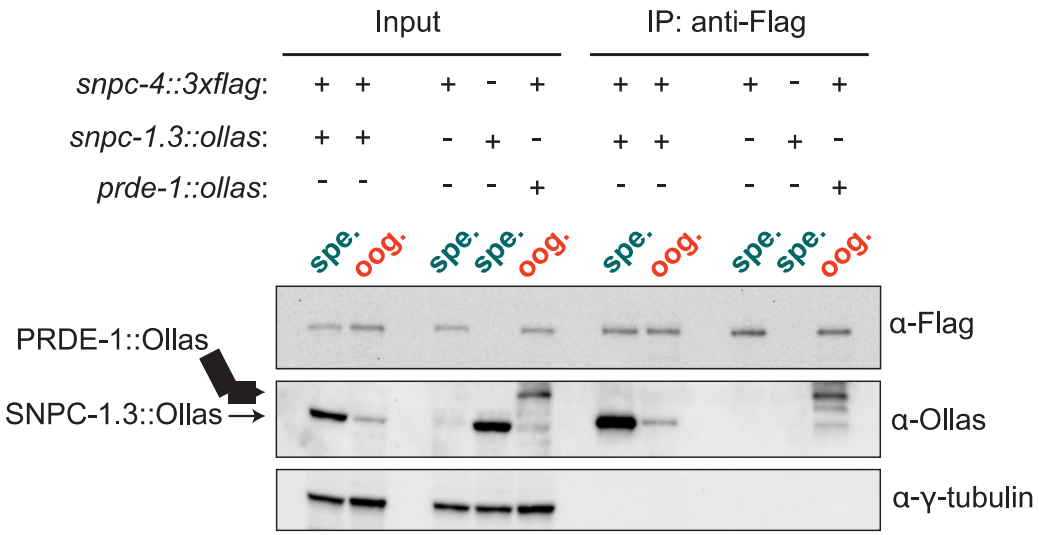

Figure 2.3. SNPC-4 interacts with SNPC-1.3.

Anti-Flag immunoprecipitation of SNPC-4::3xFlag and western blot for SNPC-1.3::Ollas during spermatogenesis (spe.) and oogenesis (oog.). PRDE-1::Ollas was used as a positive control for interaction with SNPC-4::3xFlag (Kasper et al., 2014). **γ**-tubulin was used as the loading control.

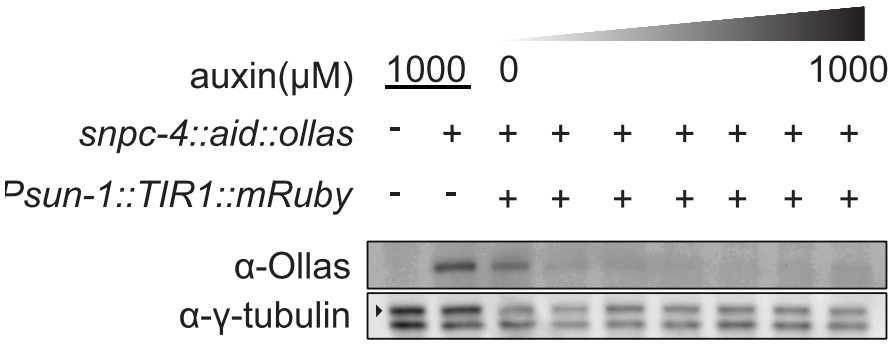

Figure 2.4. SNPC-4::AID is substantially degraded at 250 μM auxin in the germline.
Western blot of SNPC-4::AID::Ollas in worms placed on various auxin concentrations (0, 25, 50, 100, 250, 500, 1000 μM) for 1 h. The germline promoter *Psun-1* was used to drive expression of the *A. thaliana* TIR1. **γ**-tubulin is the loading control.

	Peptide counts (# of detections)	
Proteins	*fem-1*	*him-8*
SNPC-1.1	37 (2/2)	36 (2/2)
SNPC-1.2	42 (1/2)	4 (1/2)
SNPC-1.3	0 (2/2)	7 (2/2)
SNPC-1.5	15 (1/2)	0 (2/2)
SNPC-3.1	13 (2/2)	48 (2/2)
SNPC-3.2	13 (2/2)	48 (2/2)
SNPC-3.3	17 (1/2)	4 (1/2)
SNPC-3.4	12 (2/2)	14 (2/2)
PRDE-1	8 (2/2)	2 (2/2)
TOFU-5	36 (2/2)	43 (2/2)

Figure 2.5. Peptide counts of SNAPc homologs and piRNA biogenesis proteins.

Peptide counts are from first biological replicate in immunopurified samples of SNPC-4::3xFlag identified by mass spectrometry of immunopurified SNPC-4::3xFlag from *fem-1(-)* and *him-8(-)* strains.

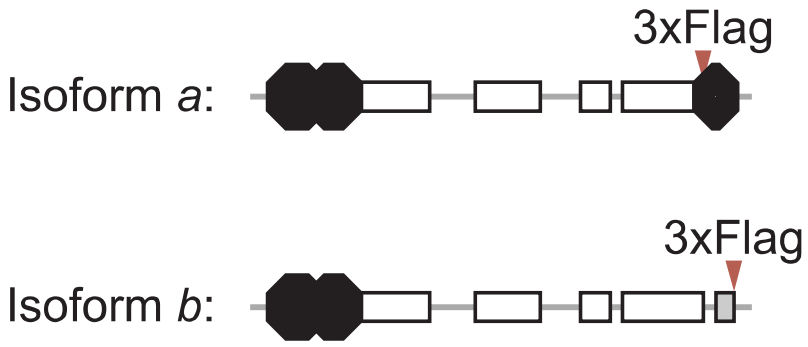

Figure 2.6. Schematic of *snpc-1.3* locus showing the location of the two different isoform-specific tags.

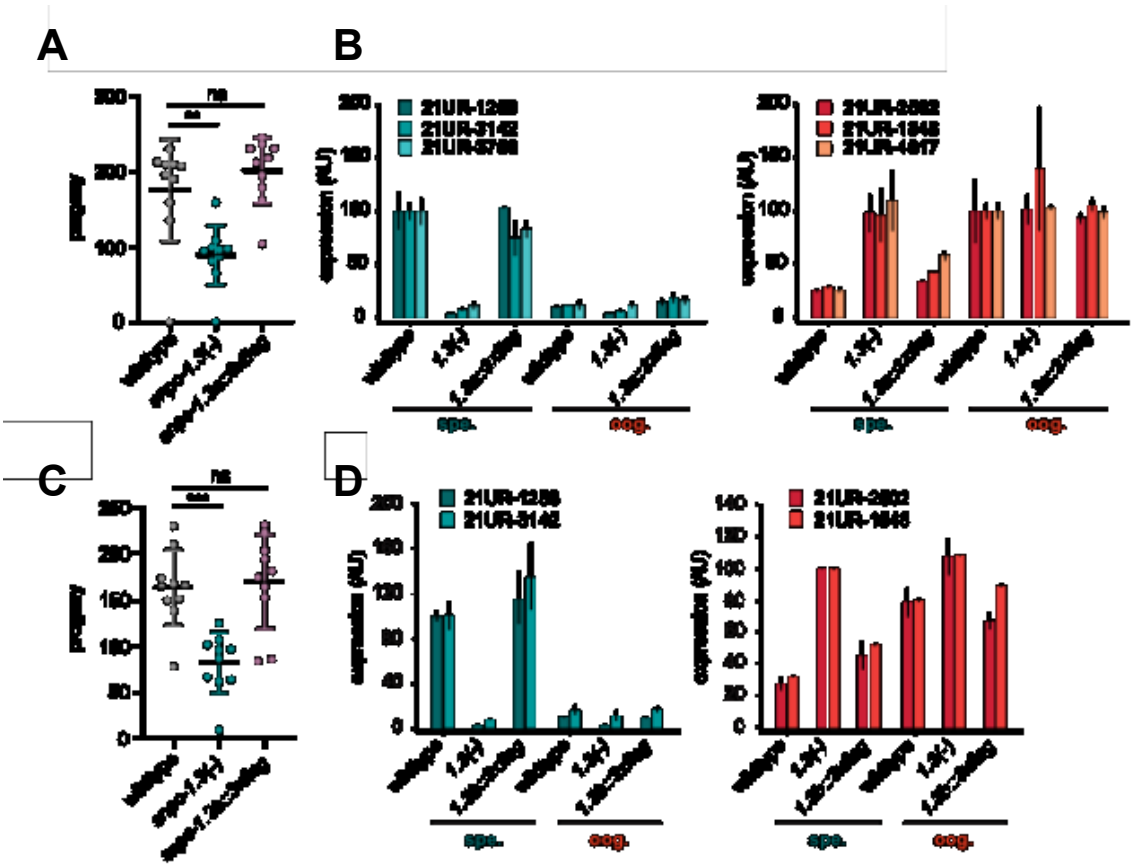

Figure 2.7. Phenotyping of snpc-1.3a::3xflag and snpc-1.3b::3xflag

(A) *snpc-1.3a::3xflag* strain has wild-type fertility at 25°C. Black bars indicate mean ± SD of n = 10 worms (wildtype versus *snpc-1.3(-)* mutant **p ≤ 0.005, Welch's t-test). (B) *snpc-1.3a::3xflag* strain has wild-type levels of male and female piRNAs during spermatogenesis and oogenesis. *1.3(-)* denotes *snpc-1.3(-)*. *1.3a::3xflag* denotes *snpc-1.3a::3xflag.* Error bars: ± SD from two technical replicates. (C) s*npc-1.3b::3xflag* strain has wild-type fertility at 25°C. Black bars indicate mean ± SD of n = 10 worms (wildtype versus *snpc-1.3(-)* mutant ***p ≤ 0.0001, Welch's t-test). (D) *snpc-1.3b::3xflag* strain has wild-type levels of male and female piRNAs during spermatogenesis and oogenesis. *1.3(-)* denotes *snpc-1.3(-). 1.3b::3xflag* denotes *snpc-1.3b::3xflag.* Error bars: ± SD from two technical replicates.

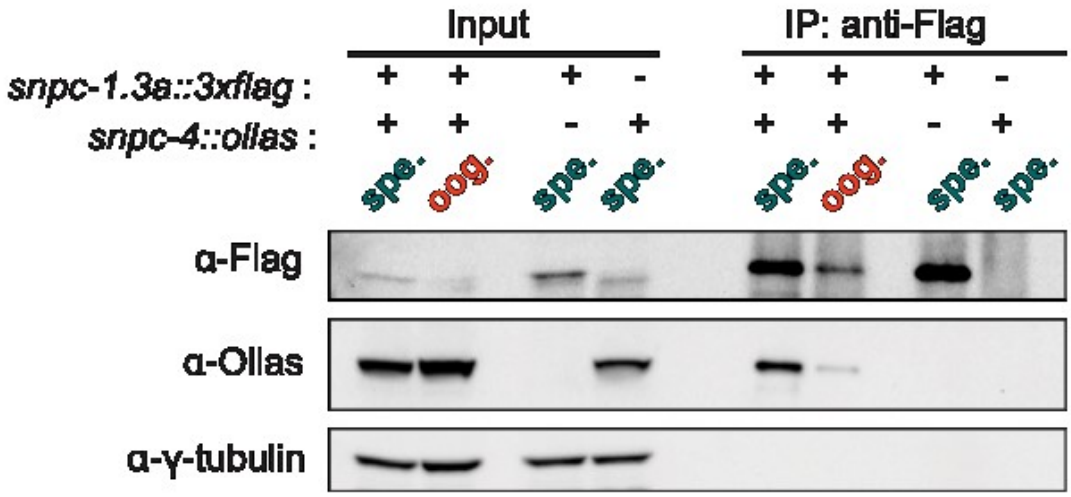

Figure 2.8. SNPC-1.3 interacts with SNPC-4.

Reciprocal immunoprecipitation of Figure 1D. Anti-Flag immunoprecipitation of SNPC-1.3::3xFlag and western blot of SNPC-4::Ollas during spermatogenesis and oogenesis. **γ**-tubulin is the loading control.

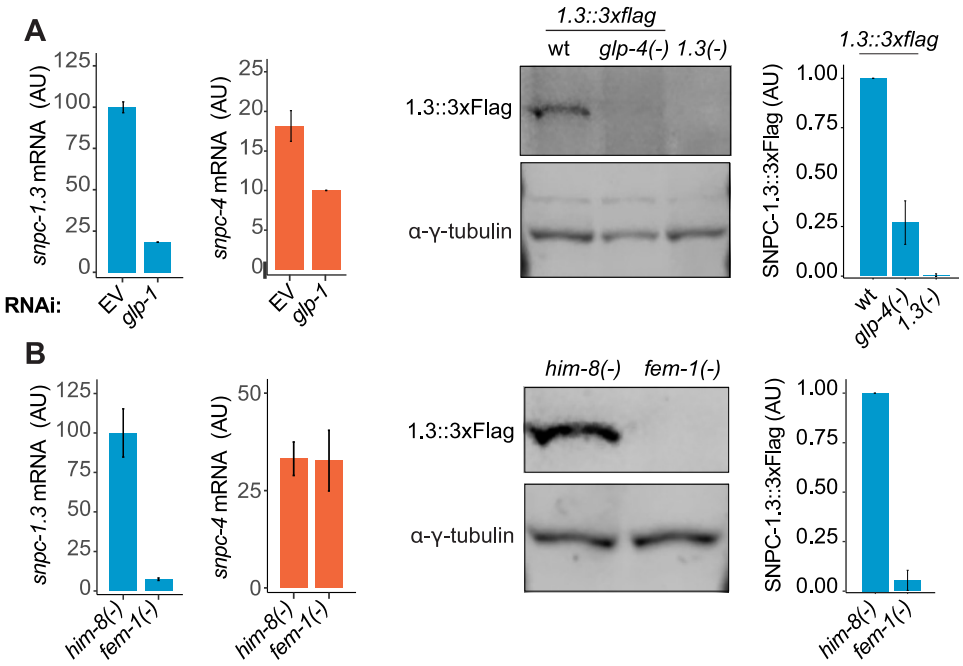

Figure 2.9. SNPC-1.3 is enriched in the male germline.

(A) SNPC-1.3 is predominantly germline-expressed. (Left) *snpc-1.3* mRNA expression is reduced upon RNAi-mediated knockdown of *glp-1* during early spermatogenesis (36 h). The housekeeping gene *eft-2* was used for normalization. Error bars: ± SD of two technical replicates. (Right) Western blot and quantification of SNPC-1.3::3xFlag in wildtype, *glp-4(-)*, and *snpc-1.3(-)* (no-Flag control) during spermatogenesis. Error bars: ± SD of two biological replicates. **γ**-tubulin was used as the loading control. (B) SNPC-1.3 is more highly expressed in males. (Left) *snpc-1.3* mRNA expression is dramatically enriched in *him-8(-)* males over *fem-1(-)* females during spermatogenesis, whereas *snpc-4* mRNA expression shows no specific enrichment. *eft-2* was used for normalization. Error bars: ± SD of two technical replicates. (Right) Western blot and quantification of SNPC-1.3::3xFlag in *him-8(-)* and *fem-1(-)*. Error bars: ± SD of two biological replicates. **γ**-tubulin was used as the loading control.

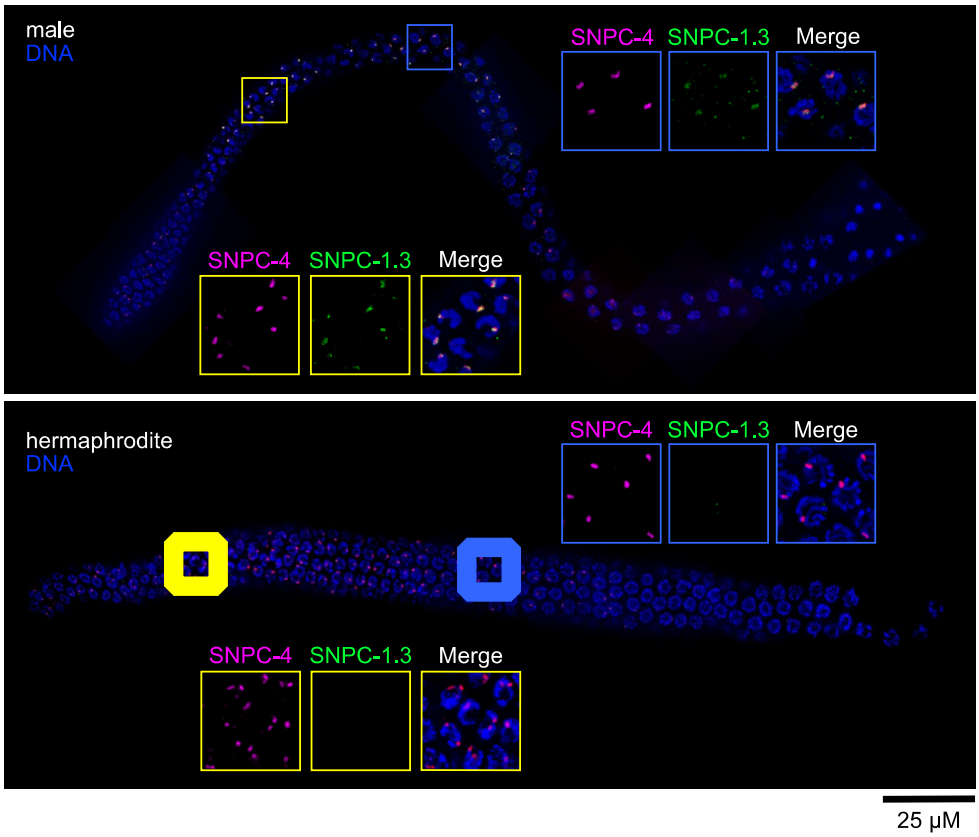

Figure 2.10. SNPC-1.3 co-localizes with SNPC-4 in the male germline.

Dissected adult male (top) and hermaphrodite (bottom) germlines stained for DNA, SNPC-4::3xFlag (magenta) and SNPC-1.3::Ollas (green) in a N2 background. Yellow insets: transition zone. Blue insets: pachytene. Representative image of three biological replicates is shown (male, n = 21,18,15 and hermaphrodite, n = 18,10,10 for hermaphrodite). Scale bar, 25 µm.

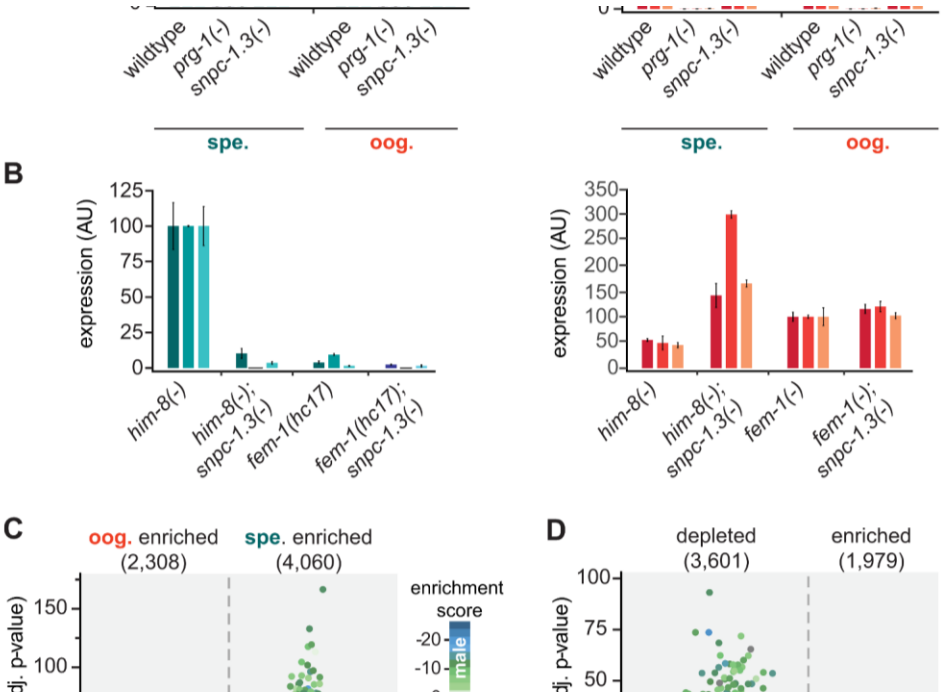

Figure 2.11. snpc-1.3 is required for male piRNA expression (spe.) but is dispensable for female piRNA expression during oogenesis (oog.).
(A) Taqman qPCR and quantification of representative male (left) and female (right) piRNAs at spermatogenic and oogenic time points normalized to U18. Error bars: ± SD of two technical replicates. (B) *him-8(-)*; *snpc-1.3(-)* mutant males exhibit severely impaired male piRNA expression and enhanced female piRNA expression. *snpc-1.3* is not required for male or female piRNA expression in *fem-1(-)* females. Error bars: ± SD from two technical replicates.

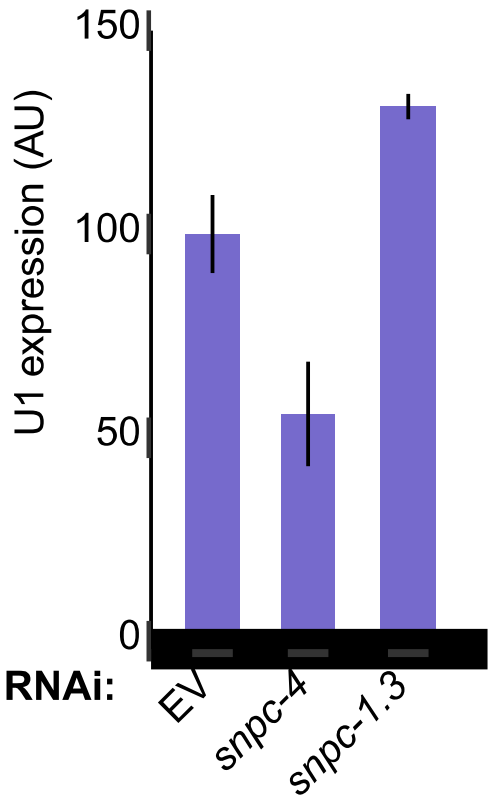

Figure 2.12. SNPC-1.3 is not responsible for U1 snRNA transcription.

U1 snRNA levels normalized to *eft-2* upon knockdown of *snpc-4* (control) and *snpc-1.3*. Knockdown of *snpc-4*, but not of *snpc-1.3*, results in a reduction in U1 levels. Error bars: ± SD of two technical replicates.

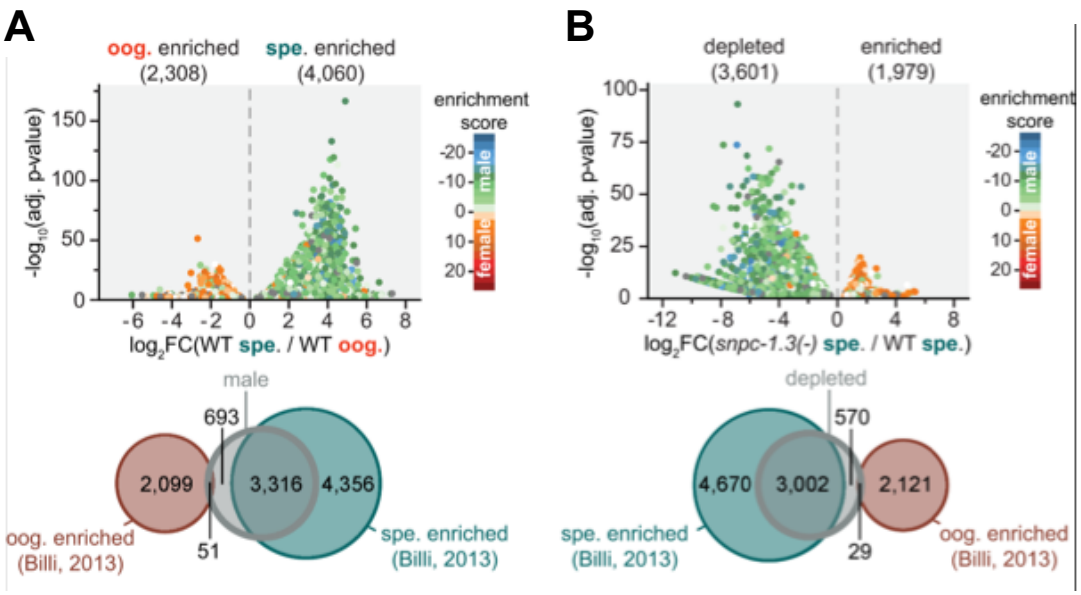

Figure 2.13. snpc-1.3 is required for male piRNA expression (spe.) but is dispensable for female piRNA expression during oogenesis (oog.).
(A) piRNAs are differentially expressed during spermatogenesis (spe.) and oogenesis (oog.) in wild-type worms. Volcano plot showing piRNAs with ≥1.2-fold change and FDR of ≤ 0.05 in 48 h (spe.) versus 72 h (oog.). piRNAs are colored according to male and female enrichment scores from Billi et al., 2013. (Bottom) Overlap of male piRNAs (spe.) in wildtype at 48 h with spermatogenesis-enriched and oogenesis-enriched piRNAs defined in Billi et al., 2013. (B) piRNAs depleted in *snpc-1.3(-)* comprise mostly of male piRNAs. (Top) Volcano plot shows piRNAs with ≥1.2-fold change and FDR ≤ 0.05 in *snpc-1.3(-)* mutant versus wildtype during spermatogenesis (spe.). piRNAs are colored according to male and female enrichment scores from Billi et al., 2013. (Bottom) Overlap of *snpc-1.3*-dependent piRNAs with spermatogenesis- and oogenesis- enriched piRNAs defined in Billi et al., 2013.

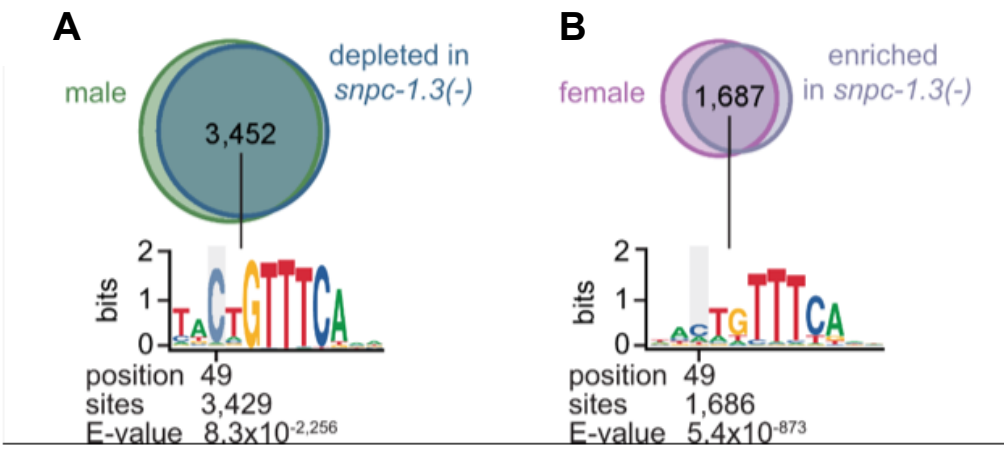

Figure 2.14. snpc-1.3 is required for male piRNA expression (spe.) but is dispensable for female piRNA expression during oogenesis (oog.).
(A) Male piRNAs that are depleted in *snpc-1.3(-)* have a conserved upstream motif with a strong 5' C bias. (Top) Overlap of *snpc-1.3*-dependent piRNAs with male piRNAs shown in Figure 3C. (Bottom) Logo plot displays conserved motif upstream of each piRNA. Median position of the C-nucleotide of the identified motif, number of piRNAs, and associated E-value are listed. (B) Female piRNAs are upregulated in *snpc-1.3(-)* mutants during spermatogenesis. (Top) Overlap of piRNAs upregulated at 72 h (oog.) with piRNAs enriched in *snpc-1.3(-)* at 48 h (spe.). (Bottom) Logo plot displays conserved motif upstream of each piRNA. Median position of the C-nucleotide of the identified motif, number of piRNAs, and associated E-value are listed.

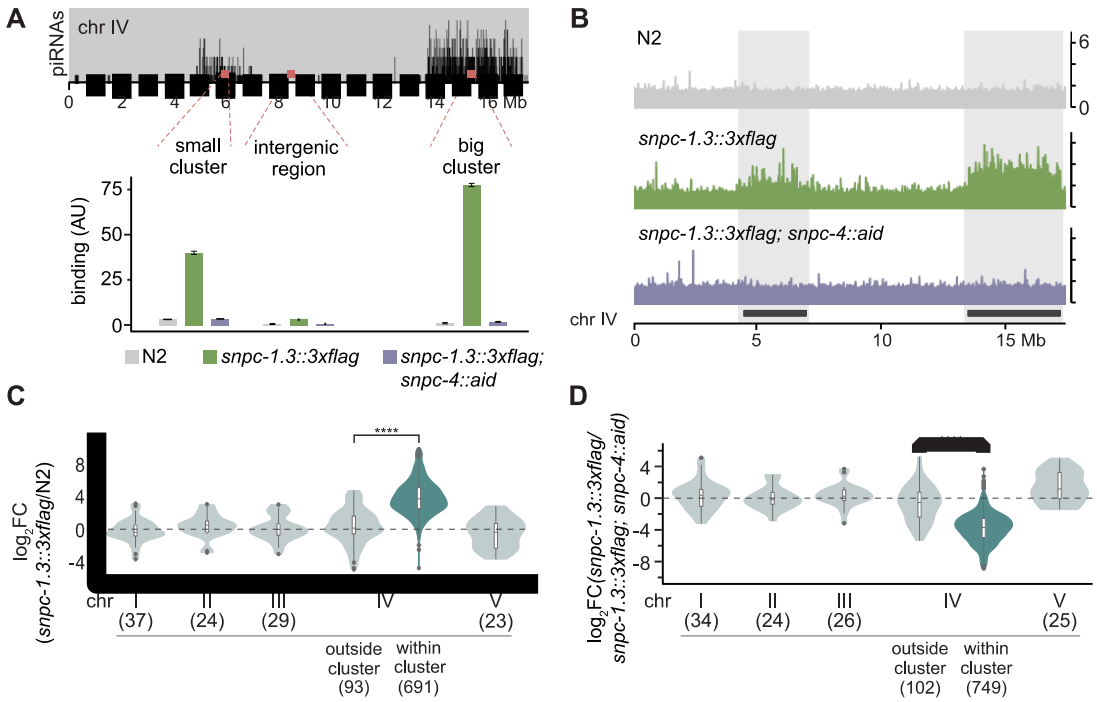

Figure 2.15. SNPC-1.3 binds male piRNA loci in a SNPC-4-dependent manner.

(A) SNPC-1.3 binding at the piRNA clusters requires SNPC-4. SNPC-1.3::3xFlag binding normalized to input (mean ± SD of two technical replicates) on chromosome IV by ChIP-qPCR in N2, *snpc-1.3::3xflag*, and *snpc-1.3::3xflag; snpc-4::aid::ollas*, which undergoes TIR-1-mediated degradation by auxin (*snpc-4::aid*). Top panel depicts the density of piRNAs on chromosome IV with piRNAs found in the small (4.5–7 Mb) and big (13.5–17.2 Mb) cluster. (B) SNPC-1.3 binding profiles across chromosome IV in N2, *snpc-1.3::3xflag*, and *snpc-1.3::3xflag*; *snpc-4::aid*. The locations of the two piRNAs clusters are highlighted. (C) SNPC-1.3 binding is enriched at piRNA clusters on chromosome IV. SNPC-1.3-bound regions are enriched within piRNA clusters compared to regions outside of the piRNA clusters on chromosome IV (**** $p \leq 0.0001$, Wilcoxon rank sum test). The number of bins analyzed is listed in parentheses. (D) SNPC-1.3 enrichment at piRNA clusters is dependent on SNPC-4. SNPC-1.3-bound regions within piRNA clusters are depleted compared to regions outside of the piRNA clusters on chromosome IV upon loss of SNPC-4 (****$p \leq 0.0001$, Wilcoxon rank sum test). The number of bins analyzed is listed in parentheses.

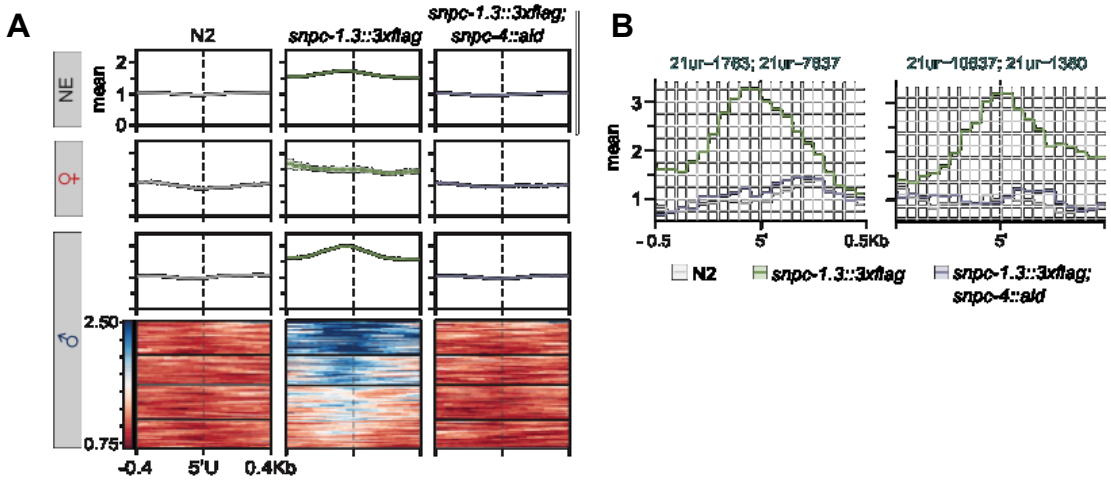

Figure 2.16. SNPC-1.3 binds male piRNA loci in a SNPC-4-dependent manner.

(A) Distribution of SNPC-1.3 reads (mean density ± standard error) around the 5' nucleotide of mature piRNAs at the piRNA clusters. To resolve SNPC-1.3 binding between male and female piRNAs despite the high density of piRNAs, we selected 1 kb bins with all male (100), female (19), or non-enriched (279) piRNAs. Heat maps represent ChIP signal in 1 kb bins around the 5' nucleotide of all 100 mature male piRNAs, ranked according to SNPC-1.3 signal. (B) Examples of SNPC-1.3 binding at two regions containing two male piRNA loci. Regions are anchored on the 5' nucleotide of each mature male piRNA and show mean read density ± standard error.

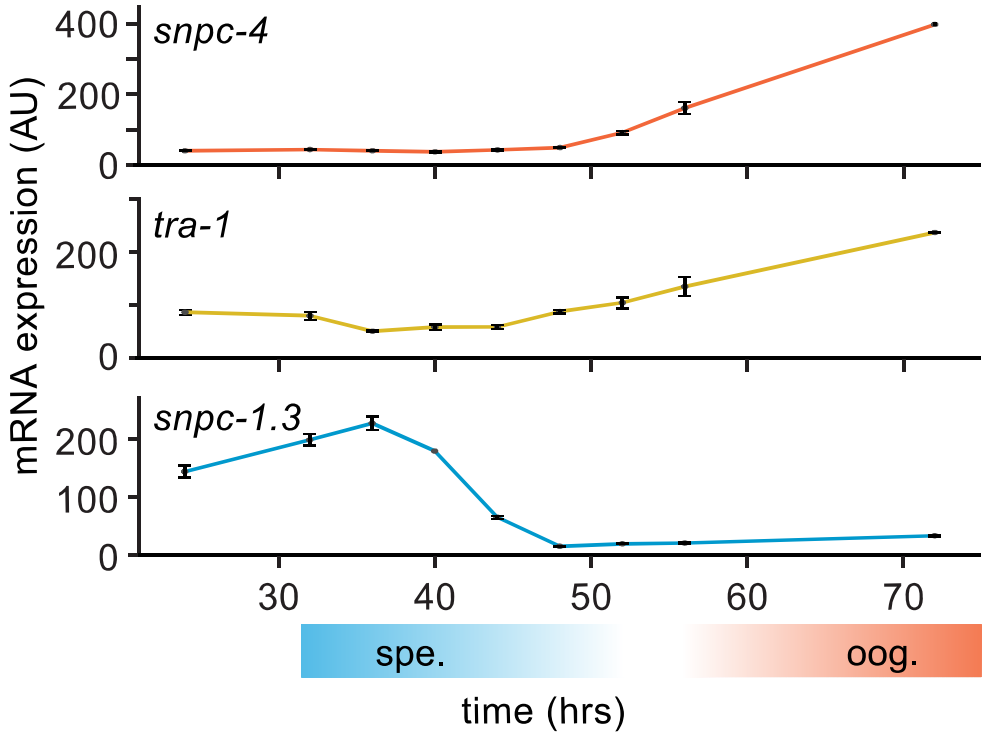

Figure 2.17. *snpc-1.3* mRNA levels peak during early spermatogenesis (spe.) while *tra-1* mRNA levels are highest during oogenesis (oog.).
qRT-PCR and quantification of *snpc-1.3, snpc-4,* and *tra-1* mRNA normalized to *eft-2* mRNA across hermaphrodite development. Time zero corresponds to the time when synchronized L1s were plated. Error bars: ± SD of two technical replicates.

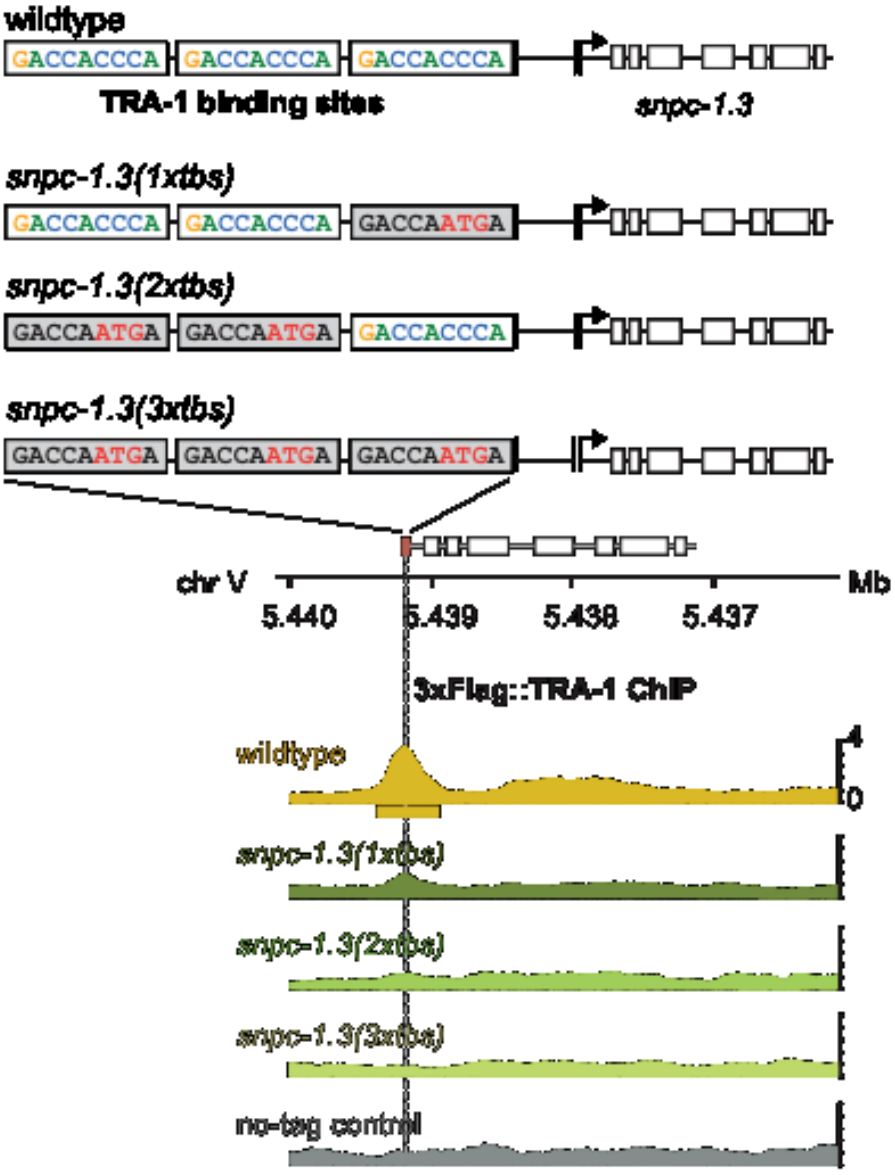

Figure 2.18. TRA-1 binds to the *snpc-1.3* promoter.

Schematic of the three TRA-1 binding sites upstream of the *snpc-1.3* locus in wildtype (top). Site-specific mutations shown in red were made in one, two, or three of the TRA-1 binding sites (grey denotes the mutated motifs). (Bottom) TRA-1 binding is reduced in TRA-1 binding site mutants assayed by TRA-1 ChIP-seq.

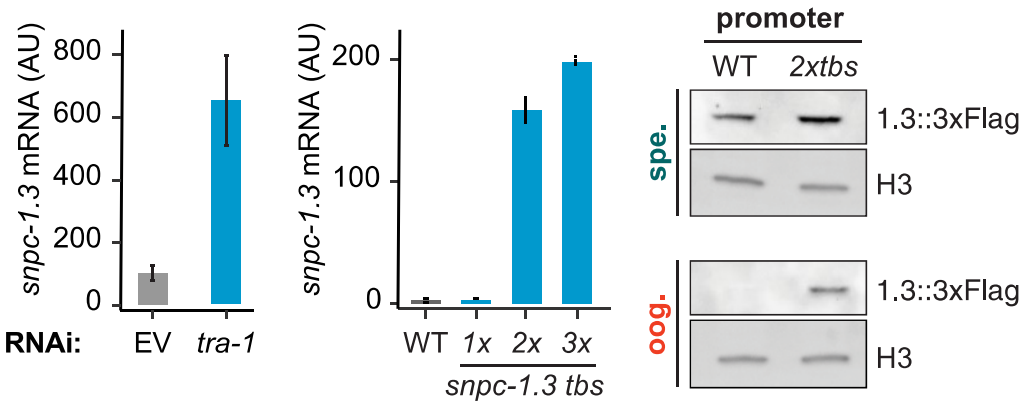

Figure 2.19. TRA-1 represses *snpc-1.3* mRNA expression during oogenesis.

(Left) *snpc-1.3* mRNA expression is drastically upregulated upon RNAi-mediated knockdown of *tra-1* and (middle) in strains bearing mutations in two (*2xtbs*) or three (*3xtbs*) TRA-1 binding sites. Error bars indicate ± SD from two technical replicates. (Right) Western blot of SNCP-1.3::3xFlag expression driven under the wild-type and *2xtbs* mutant promoter during spermatogenesis (spe.) (top) and oogenesis (oog.) (bottom). H3 was used as the loading control.

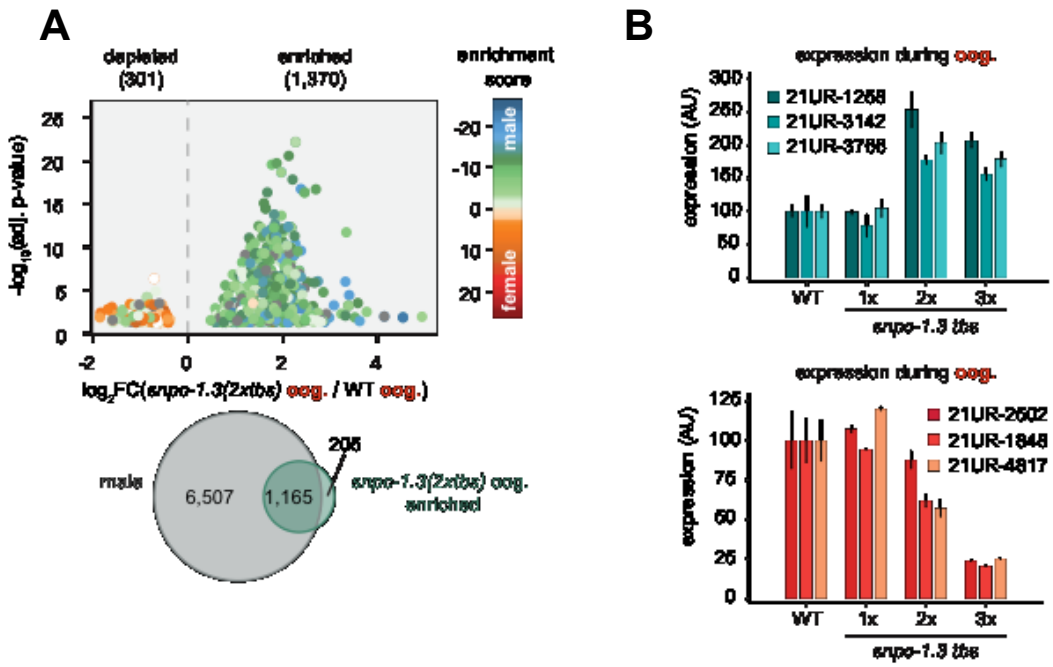

Figure 2.20. TRA-1 represses *snpc-1.3* and male piRNAs expression during oogenesis.

(A) A subset of male piRNAs are ectopically expressed during oogenesis in *snpc-1.3(2xtbs)* mutants. (Top) Volcano plot showing differential piRNA expression between *snpc-1.3(2xtbs)* mutants versus wildtype during oogenesis (oog.). piRNAs are colored by enrichment scores from Billi et al., 2013. (Bottom) Overlap of male piRNAs defined in Figure 3C with upregulated piRNAs in *snpc-1.3(2xtbs)* mutants. (B) Mutations at two (*2xtbs*) or three (*3xtbs*) TRA-1 binding sites enhance male piRNA expression (top) but attenuate female piRNA expression (bottom) during oogenesis. Error bars indicate ± SD from two technical replicates.

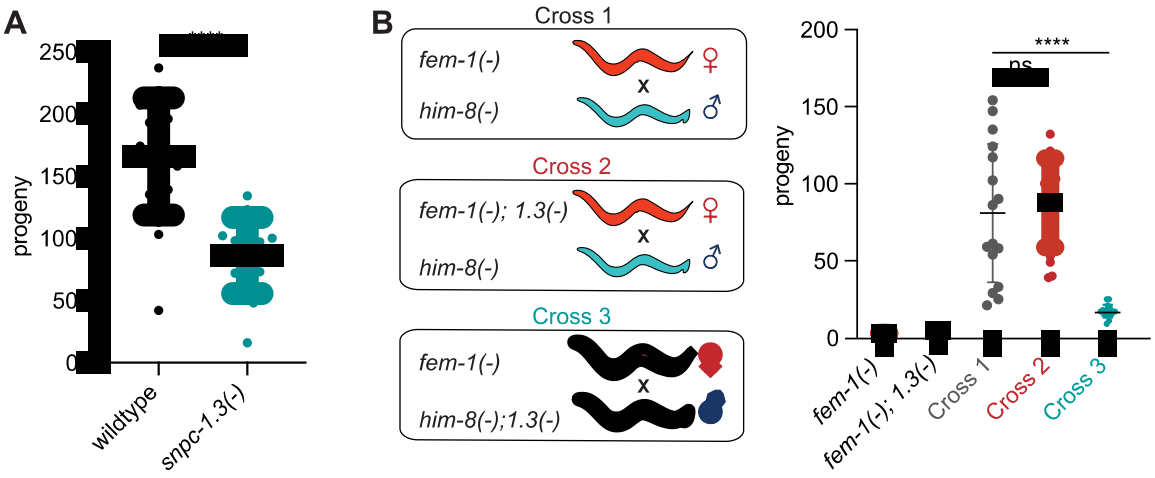

Figure 2.21. SNPC-1.3 is critical for male fertility.

(A) *snpc-1.3(-)* hermaphrodites exhibit sterility at 25°C. Circles correspond to the number of viable progeny from singled hermaphrodites (n = 16). Black bars indicate mean ± SD. Statistical significance was assessed using Welch's t-test (****p ≤ 0.0001). (B) *snpc-1.3* promotes male fertility but is dispensable for female fertility. (Left) Diagram illustrates crosses between strains for mating assays (*1.3(-)* denotes *snpc-1.3(-)*). (Right) *snpc-1.3(-)*; *him-8(-)* males crossed to *fem-1(-)* females show severe fertility defects (Cross 3). *snpc-1.3; fem-1(-)* females crossed to *him-8(-)* males (Cross 2) show equivalent fertility similar to *fem-1(-)* females crossed to *him-8(-)* males (Cross 1). Circles correspond to the number of viable progeny from cross (n = 16). Black bars indicate mean ± SD. Statistical significance was assessed using Welch's t-test (ns: not significant; ****p ≤ 0.0001).

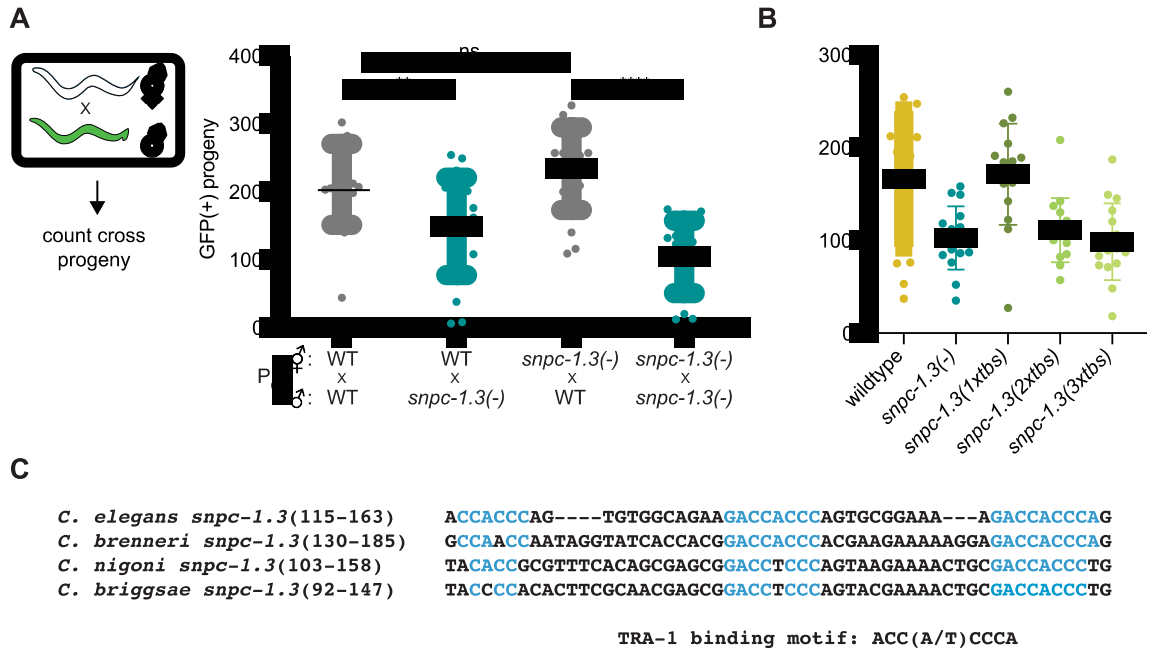

Figure 2.22. SNPC-1.3 is critical for male fertility.

(A) *snpc-1.3* is required in males, but not females, to promote fertility. (Left) Diagram illustrates crosses between strains for mating assays. Green worms represent *Pcol-19::gfp* males. (Right) *snpc-1.3(-); Pcol-19::gfp* males show diminished ability to rescue wild-type or *snpc-1.3(-)* hermaphrodites when compared to *Pcol-19::gfp* males. Circles correspond to the brood size of viable progeny from each mating (n = 16). Black bars indicate mean ± SD. Statistical significance was assessed using Welch's t-test (ns: not significant;**p ≤ 0.005; ****p ≤ 0.0005). (B) *snpc-1.3(2xtbs)* and *snpc-1.3(3xtbs)* mutant hermaphrodites have decreased fertility at 25°C. Black bars indicate mean ± SD (n = 16). Statistical significance was assessed using Welch's t-test (****p ≤ 0.0001). (C) Multiple TRA-1 binding sites are conserved in *C. elegans, C. brenneri, C. briggsae,* and *C. nigoni.* Sequence alignment of *snpc-1.3* homologs in other nematode species. Blue indicates sequences in TRA-1 binding motifs.

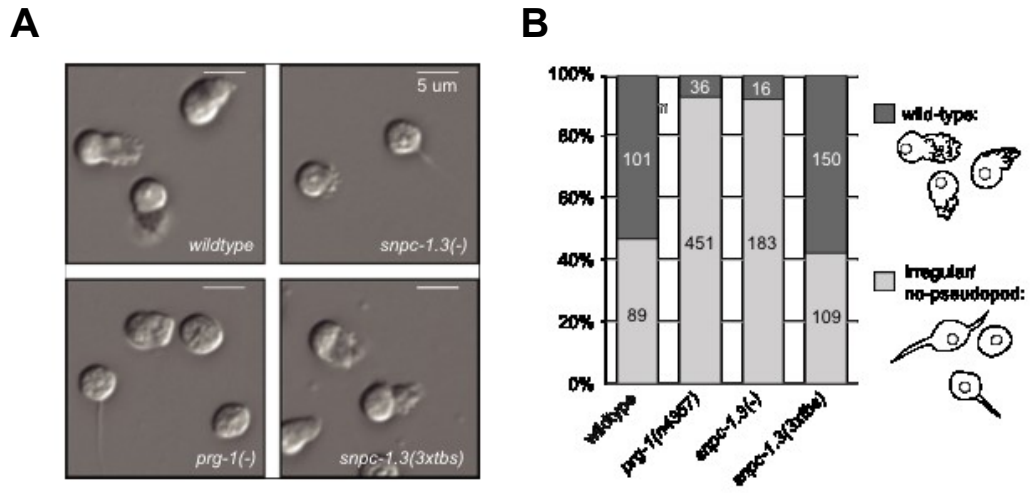

Figure 2.23. SNPC-1.3 is critical for male fertility.

(A) *snpc-1.3(-)* spermatids exhibit severe morphological defects. Images of pronase-treated sperm of wild-type, *prg-1(-), snpc-1.3(-)*, and *snpc-1.3(2xtbs)* males. (B) *snpc-1.3(-)* spermatids exhibit severe sperm maturation defects.

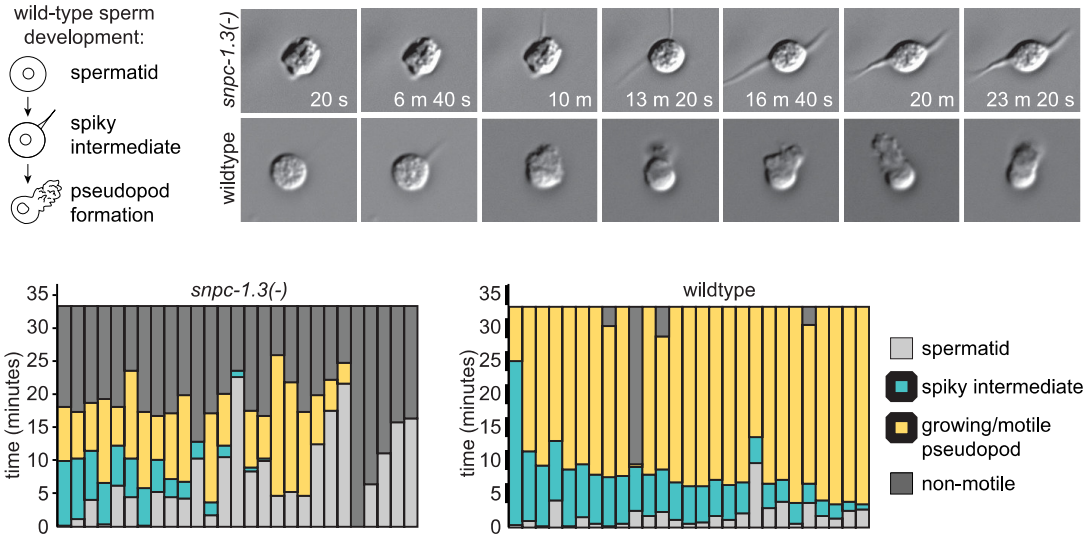

Figure 2.24. Graphical display of individual sperm tracked over time after pronase treatment.

(Top) Images depicted at 3 min intervals of a sperm undergoing activation and maturation. Imaging of spermatid commenced ~3 min after pronase treatment. (Bottom)

2.6 Discussion

Our data indicate that *C. elegans* SNPC-1.3, a human SNAPC1 ortholog, functions as a male piRNA transcription factor. SNPC-1.3 interacts with SNPC-4 in foci in male germ cell nuclei and, by preferentially binding male piRNA promoters, is critical for their expression. SNPC-1.3 expression, reflecting the developmental profile of male piRNAs, is highest during spermatogenesis. I demonstrate that the *snpc-1.3* locus itself is regulated by the sex determination regulator, TRA-1. During spermatogenesis, *tra-1* expression is low, and *snpc-1.3* and other male-promoting genes are licensed for expression. In contrast, *tra-1* expression is upregulated during oogenesis and TRA-1 binds the *snpc-1.3* promoter to repress its transcription, leading to the expression of female over male piRNAs . I propose that SNPC-1.3, via its interaction with SNPC-4, can direct the specificity of the core piRNA complex preferentially to male piRNA loci.

2.6.1 How is the expression of male and female piRNAs coordinated?

Given its role as a putative male piRNA transcription factor, I expected that deletion of *snpc-1.3* would result in loss of male piRNAs with no consequences to the expression of female piRNAs. However, loss of *snpc-1.3* also results in increased female piRNA expression during spermatogenesis (Figure 2.3), whereas ectopic overexpression of *snpc-1.3* during oogenesis leads to decreased female piRNA levels. Taken together, our findings suggest that transcription of male and female piRNAs is not completely separable from each other and that the balance in expression of the two piRNA subclasses may be dictated by the allocation of shared core transcription factors such as SNPC-4.

Similar to multiple gene classes activated by general transcription factors (Levine et al., 2014), I speculate that male and female promoters compete for access to a limited pool of the core transcription complex, which includes SNPC-4, PRDE-1, TOFU-4, and TOFU-5. Therefore, we propose a model in which the expression and binding of SNPC-1.3 to core piRNA factors serves to “sequester” the core complex away from female promoters. Mechanistically, I posit that the core piRNA transcription complex is specified to female promoters, and that only upon association with SNPC-1.3 is the core machinery directed to male promoters. I predict that when SNPC-1.3 is absent, more SNPC-4 and other previously identified cofactors are available to transcribe female piRNAs. Conversely, overexpression of SNPC-1.3 leads to the disproportionate recruitment of the core machinery to male promoters, leading to the indirect downregulation of female piRNAs. By controlling male piRNA expression, SNPC-1.3 is crucial for maintaining the balance between male and female piRNA levels across development.

While the default specification of the core complex to female promoters presents perhaps the most parsimonious explanation underlying male and female piRNA expression, I cannot exclude the possibility that an additional female-specific *trans*-acting factor may direct the core piRNA complex to female promoters. If true, I speculate that the developmental expression of such a factor (low during spermatogenesis and high during oogenesis), coupled with the developmental expression of SNPC-1.3, would coordinate the differential expression of male and female piRNAs. During spermatogenesis, SNPC-1.3 is

more highly expressed such that the core machinery would primarily be directed to male promoters. In contrast, during oogenesis, SNPC-1.3 expression is low, concomitant with elevated expression of a female factor to license transcription of female piRNAs. This model, where both factors are present during both spermatogenesis and oogenesis but in different ratios, would also be consistent with our piRNA expression analysis in *snpc-1.3* loss-of-function and overexpression mutants.

2.6.2 The piRNA pathway co-opts snRNA biogenesis machinery.

My work adds to a growing body of evidence that snRNA machinery has been hijacked at multiple stages in *C. elegans* piRNA biogenesis, including transcription (Kasper et al., 2014; Weng et al., 2019) and termination (Beltran et al., 2019). Investigating potential parallels between snRNA and piRNA biogenesis may provide useful clues into the role of SNPC-1.3 in the piRNA complex.

The minimal snRNA SNAP complex consists of a 1:1:1 heterotrimer of the subunits SNAPC4, SNAPC1, and SNAPC3 in humans and SNAP190, SNAP43, and SNAP50 in flies (Henry et al., 1998; Hung and Stumph, 2011; Li et al., 2004; Ma and Hernandez, 2002; Mittal et al., 1999) (Figure 2.1B). *In vitro* studies have shown that the trimer must assemble before the complex is able to bind DNA. Similarly, our data show SNPC-1.3 requires SNPC-4 to bind at the piRNA clusters, although we cannot formally rule out that loss of SNPC-4 only affects the stability of SNPC-1.3, rather than directly recruiting SNPC-1.3 to piRNA promoters. I speculate the piRNA complex is assembled in a similar fashion to the snRNA complex. Based on this model, we expect that SNPC-4 binding at male piRNA loci

is abolished in a *snpc-1.3* mutant. However, conclusive evidence that SNPC-4 binding at male piRNA promoters requires SNPC-1.3 is still lacking. Due to the highly clustered nature of *C. elegans* piRNAs, we anticipate that detecting differences in SNPC-4 binding between male and female piRNAs in *snpc-1.3(-)* mutants may not be possible with traditional ChIP-seq methods, and may require application of higher resolution techniques.

Given that piRNAs have co-opted *trans*-acting factors from snRNA biogenesis (Kasper et al., 2014), it would not be surprising if piRNAs also co-evolved *cis*-regulatory elements for transcription factor binding from snRNA loci. Recently, Beltran et al. (2019) identified similarity between the 3' end of PSEs of snRNA promoters and the 8 nt piRNA core motif in nematodes. In addition, Pol II and Pol III transcription from snRNA promoters share a common PSE, but are distinguished by the presence of other unique motifs (Hung and Stumph, 2011). Correspondingly, the canonical Type I and less abundant Type II piRNAs can be discriminated by the presence or absence of the 8 nt core motif, respectively. Factors such as TOFU-4 and TOFU-5 function in both Type I and II piRNA expression, whereas PRDE-1 is only required for Type I piRNAs (Kasper et al., 2014; Weng et al., 2019). Altogether, these observations highlight the importance of *cis*-regulatory elements in specifying the expression of snRNAs and piRNA classes. In addition to enrichment of cytosine at the 5' position in the male core motif (Billi et al., 2013), we hypothesize that as-yet unidentified motifs may further discriminate male from female piRNA promoters. While we observed SNPC-1.3 binding to be enriched upstream of male piRNA loci, we cannot definitively

conclude that SNPC-1.3 binds to the male-specific core motif, given the limitations of conventional ChIP-seq in resolving the SNPC-1.3 footprint. Identifying the factors that specifically bind the 8 nt core motif and other potential *cis*-regulatory elements important for sex-specific piRNA expression will require further investigation.

2.6.3 What are the functions of male piRNAs in *C. elegans*?

My data suggest that SNPC-1.3 is essential for proper spermiogenesis (Figure 2.6). We hypothesize the global loss of male piRNAs in a *snpc-1.3(-)* mutant is responsible for the higher incidence of spermiogenesis arrest and subsequent loss in fertility, although it is possible that SNPC-1.3 may have other or additional effects on male fertility. Characterization of *prg-1(-)* mutants during spermiogenesis agree with our findings that loss of piRNAs in the male germline leads to acute defects directly responsible for fertility (Wang and Reinke, 2008). Since the initial discovery of piRNA function in the targeting and silencing of transposons in *Drosophila* (Vagin et al., 2006; Brennecke et al., 2007), analyses in other systems have revealed that piRNAs have acquired neofunctions at later points along the evolutionary timescale (Ozata et al., 2019).

While it is estimated that as much as 45% of the human genome encodes for transposable elements (Lander et al., 2001), only 12% of *C. elegans* genome encodes such elements. Furthermore, nearly all of these regions are inactive in *C. elegans* (Bessereau, 2006). In contrast to *Drosophila* piRNAs that target and silence transposons with perfect complementarity (Brennecke et al., 2007), *C. elegans* piRNAs are thought to bind a broad range of endogenously expressed

transcripts by partial complementarity (Ashe et al., 2012; Shen et al., 2018; Zhang et al., 2018). Together, these findings suggest that worm piRNAs function in capacities distinct from canonical transposon silencing. While a recent methodology used cross-linking, ligation, and sequencing of piRNA:target hybrids (CLASH) to determine that female piRNAs engage with almost every germline transcript (Shen et al., 2018), how female piRNAs select their targets has yet to be examined. Like piRNAs characterized in the female germline, male piRNAs may be interfacing with a broad range of targets to regulate gene expression for proper spermatogenesis. Loss of *prg-1* in males causes the downregulation of a subset of spermatogenesis-specific genes (Wang and Reinke, 2008), suggesting male piRNAs serve a protective function for spermatogenic processes. The characterization of the *in vivo* landscape of male piRNA target selection using CLASH may provide insights into piRNA function during spermatogenesis.

2.6.4 Why are male piRNAs restricted from the female germline?

Sperm and oocytes pass epigenetic information such as noncoding RNAs to the next generation (Hammoud et al., 2014; Brykczynska et al., 2010; Tabuchi et al., 2018; Kaneshiro et al., 2019). Recent studies show maternal piRNAs trigger the production of endo-siRNAs, called 22G-RNAs for their 5' bias for guanine and 22 nt length, to transmit an epigenetic memory of foreign versus endogenous elements to the next generation (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). We predict that misexpression of male piRNAs during oogenesis may perturb the native pool of female piRNAs necessary for appropriate recognition of self versus non-self elements. This may explain the decrease in fertility we

observed in multiple TRA-1 binding site mutant hermaphrodites (Figure 2.19B). As *snpc-1.3(3xtbs)* sperm do not seem to exhibit significant morphological defects (Figure 6), the fertility defects in the *snpc-1.3(3xtbs)* mutants could be due to problems arising in oogenesis. However, based on our sequencing data in *snpc-1.3(2xtbs)* mutants, we cannot distinguish whether fertility defects during oogenesis are due to upregulation of male piRNAs, downregulation of female piRNAs, a combination of the two, or misexpression of downstream endo-siRNAs triggered by piRNAs. Further study of *snpc-1.3* gain-of-function mutants in oogenesis will enhance our understanding of the physiological consequences of expressing male piRNAs in the female germline.

2.6.5 The intersection between sex determination and sex-specification of piRNA expression.

I speculate that gene duplication of the *snpc-1* family of genes occurred early during nematode evolution and allowed for the acquisition of new functions by *snpc-1* paralogs, specifically, from snRNA to piRNA biogenesis. At least two SNPC-1 paralogs are present within the distantly related nematode species, *Plectus sambesii.* Furthermore, I predict that co-opting SNPC-1 paralogs for piRNA biogenesis may have occurred in parallel with the evolution of the nematode sex determination pathway. TRA-1 is a sex determination factor that acts to repress male-promoting gene expression in female germ cells to promote female germ cell fate. While *Drosophila* sex determination utilizes different factors than *C. elegans*, further investigation into the conservation of TRA-1 shows that it is a common feature in at least the nematode lineage (Pires-daSilva and Sommer, 2004).

Additionally, just as we have shown that TRA-1 represses *snpc-1.3* in *C. elegans*, TRA-1 binding motifs GGG(A/T)GG are present in the putative upstream promoter regions of *snpc-1.3* homologs identified in *C. briggsae, C. brenneri* and *C. nigoni* (Figure 2.19C). Taken together, these analyses point to a conserved link between sex determination and piRNA biogenesis pathways among nematodes.

In summary, our work reveals that SNPC-1.3 is specified to the male germline and is essential for male piRNA expression. We have identified SNPC-1.3 as a major target of TRA-1 repression in the female germline. Future studies will likely uncover additional factors required to coordinate the proper balance of sex-specific piRNAs required for proper germline development and animal fertility.

2.7 Experimental model and subject details.

C. elegans strains were maintained at 20°C according to standard procedures (Brenner, 1974), unless otherwise stated. Bristol N2 was used as the wild-type strain. Except for RNAi and ChIP experiments, worms were fed *E. coli* strain OP50. Worms used for ChIP were fed *E. coli* strain HB101.

2.8 Materials and Methods

Generations of strains.

CRISPR/Cas9-generated strains were created as described in Paix et al., 2015 and are listed in Supplementary file 4. crRNA and repair template sequences of CRISPR-generated strains are listed in Supplementary file 4. After initial phenotyping of *snpc-1.3a::3xflag* and *snpc-1.3b::3xflag* (Figure 2.4, 2.5A–D), *snpc-1.3a::3xflag* was used for all subsequent experiments (and is referred to as *snpc-1.3::3xflag*).

RNAi assays.

Bacterial RNAi clones were grown from the Ahringer RNAi library (Kamath and Ahringer, 2003). Synchronized L1 worms were plated on HT115 bacteria expressing dsRNA targeting the gene interest or L4440 empty vector as a negative control as previously described (Timmons and Fire, 1998). All RNAi experiments were performed at 20°C unless otherwise stated.

RNA extraction, library preparation, and sequencing.

After hypochlorite preparation and hatching in M9 buffer, *snpc-4::aid::ollas* and *snpc-4::aid::ollas; Psun-1::TIR1* worms were transferred from NGM plates to plates containing 250 µM auxin 20 h before collecting L4 and gravid worms, 48 h and 72 h after plating L1 worms at 20°C, respectively. Worms were collected in TriReagent (Thermo Fisher Scientific) and subjected to three freeze-thaw cycles. Following addition of 1-bromo-3-chloropropane (BCP), the aqueous phase was then precipitated with isopropanol at -80°C for 2 h. To pellet RNA, samples were spun at 21,000 x *g* for 30 min at 4°C. After three washes in 75% ethanol, the pellet was resuspended in water.

RNA concentration and quality were measured using a TapeStation (Agilent Technologies). 16–30 nt small RNAs were size-selected from 5 µg total RNA on 17% denaturing polyacrylamide gels. Small RNAs were treated with 5’ polyphosphatase (Illumina) to reduce 5’ triphosphate groups to monophosphates to enable 5’ adapter ligation. Small RNA-sequencing libraries were prepared using the NEBNext® Multiplex Small RNA Library Prep Set for Illumina (NEB). Small RNA amplicons were size-selected on 10% polyacrylamide gels and quantified

using qRT-PCR. Samples for each developmental time point were pooled into a single flow cell and single-end, 75 nt reads were generated on a NextSeq 500 (Illumina). An average of 42.01 million reads (range 33.05–50.39 million) was obtained for each library.

Quantitative RT-PCR.

Taqman cDNA synthesis was performed as previously described (Weiser et al., 2017). Briefly, for quantification of piRNA levels, TaqMan small RNA probes were designed and synthesized by Applied Biosystems. All piRNA species assessed by qPCR were normalized to U18 small nucleolar RNA. 50 ng of total RNA was used for cDNA synthesis. cDNA was synthesized by Multiscribe Reverse Transcriptase (Applied Biosystems) using the Eppendorf Mastercycler Pro S6325 (Eppendorf). Detection of small RNAs was performed using the TaqMan Universal PCR Master Mix and No AmpErase® UNG (Applied Biosystems). For quantification of mRNA levels, cDNA was made using 500 ng of total RNA using Multiscribe Reverse Transcriptase (Applied Biosystems). For quantification of snRNA levels, cDNA was made using 250 ng of total RNA using SuperScript III Reverse Transcriptase (ThermoFisher). Assays for mRNA and snRNA levels were performed with Absolute Blue SYBR Green (ThermoFisher) and normalized to *eft-2* using CFX63 Real Time System Thermocyclers (Biorad). All qPCR primers used are listed in Supplementary file 4.

Covalent crosslinking of Dynabeads.

Protein G Dynabeads (ThermoFisher Scientific, 1003D) were coupled to monoclonal mouse anti-FLAG antibody M2 (Sigma Aldrich, F1804). After 3 washes

in 1x PBST (0.1% Tween), Dynabeads were resuspended with 1x PBST with antibody, for a final concentration of 50 µg antibody per 100 µL beads. The antibody-bead mixture was nutated for 1 h at room temperature. After 3 washes in 1x PBST and 2 washes in 0.2 M sodium borate pH 9.0, beads were nutated in 22 mM DMP (Sigma Aldrich, D8388) in 0.2 M sodium borate for 30 min at room temperature. Following 2 washes in ethanolamine buffer (0.2 M ethanolamine, 0.2 M NaCl pH 8.5), beads were nutated for 1 h at room temperature in the same buffer. Beads were placed into the same volume of ethanolamine buffer as the starting bead volume for storage at 4°C until use.

Immunoprecipitation for mass spectrometry, co-IP experiments, and expression.

For SNPC-4 IP mass spectrometry, synchronized populations of ~200,000 *him-8(e1489)* L4s and ~50,000 *fem-1(hc17)* females were grown at 25°C and collected on OP50**.** For co-IP experiments, ~500,000 L4 and ~250,000 gravid worms were grown and collected from OP50 plates. Due to low expression of SNPC-1.3 and appearance of background bands, samples examining SNPC-1.3 expression were subjected to immunoprecipitation before western blotting. For *glp-4(bn2)*, *him-8(e1489)*, *and fem-1(hc17)* temperature-shift experiments, worms were grown at 15°C before hypochlorite treatment to isolate embryos. Synchronized L1s were then transferred to 25°C. For SNPC-1.3 expression in males and females, *snpc-1.3::3xflag; him-8(e1489)* L4 worms and *snpc-1.3::3xflag; fem-1(hc17)* adult worms were collected.

Unless otherwise stated, all samples for mass spectrometry, co-IP, and western blotting used in this study were subjected to the following procedure. After three washes in M9 and one wash in water, worms were frozen and ground using the Retsch MM400 ball mill homogenizer for 2 rounds of 1 min at 30 Hz. Frozen worm powder was resuspended in 1x lysis buffer used previously (Moissiard et al., 2014) (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM EGTA, 0.1% NP-40, 10% glycerol) and protease inhibitor cocktail (Roche). After Bradford assay (ThermoFisher Scientific), lysates were normalized using lysis buffer and protease inhibitor. Benzonase (Sigma-Aldrich, E1014) was added to a final concentration of 1 µL/mL of lysate and nutated for 10 min at 4°C. After centrifugation for 10 min at 4,000 x *g*, 1 mL of supernatant was added to 50 µL of crosslinked Dynabeads and nutated for 15 min at 4°C. Samples were then washed 3 times in 1x lysis buffer with protease inhibitors before 1 h nutation in 50 µL of 2 mg/mL FLAG peptide (Sigma-Aldrich, F4799) diluted in 1x lysis buffer. Complete eluate, as well as 5% of crude lysate (after addition of benzonase), input, pellet, and post-IP samples were added to 2x Novex Tris-Glycine SDS Sample Buffer (ThermoFisher Scientific, LC2676) to 1x. Samples were then subjected to western blotting as described below.

Western blotting.

Co-IP samples and SNPC-1.3::3xFlag westerns in *snpc-1.3 tbs* mutants were run on either 8–16% or 8% Novex WedgeWell Tris-Glycine precast gels (ThermoFisher), and transferred to PVDF membrane (Millipore). Mouse anti-Flag, rat anti-Ollas, rabbit anti-gamma tubulin, and rabbit anti-H3 were used at 1:1,000,

1:8,000, 1:5,000, and 1:15,000, respectively. Anti-mouse and anti-rabbit (for tubulin) antibodies were used at 1:5,000. To blot for H3, anti-rabbit secondary was used at 1:15,000. Anti-rat antibodies were used at 1:8,000. Antibodies used were Sigma-Aldrich F1804 (mouse anti-Flag), Novus Biologicals NBP1-06713SS (rat anti-Ollas), Sigma-Aldrich T1450 (rabbit anti-gamma tubulin), and Abcam ab1791 (rabbit anti-H3), GE Healthcare NA931 (sheep anti-mouse), and Jackson Laboratories 111035045 (goat anti-rabbit). Both high sensitivity Amersham ECL Prime (GE Healthcare, RPN2232) (for SNPC-1.3 blotting) and regular sensitivity Pierce ECL (ThermoFisher, 32209) were used for exposure in a BioRad ChemiDoc Touch system.

For measuring SNPC-1.3 expression levels in various backgrounds, input (for normalization) and immunoprecipitation samples were run on 10% Novex WedgeWell Tris-Glycine precast gels (ThermoFisher). Following transfer, the membrane was dried for 20 min at room temperature. The blot was then recharged in 100% methanol for 1 min, followed by a water rinse and a wash in TBS for 2 min. Blocking was performed in LICOR Odyssey Blocking Buffer (TBS). Primary antibodies were 1:1,000 mouse anti-Flag (Sigma-Aldrich F1804) and 1:5,000 rabbit anti-gamma-tubulin (Sigma-Aldrich T1450) in LICOR Odyssey Blocking Buffer with 0.1% Tween. Washes were performed in TBST (TBS + 0.1% Tween). LICOR IRDye 800CW goat anti-mouse IgG and 680RD goat anti-rabbit IgG were used at 1:15,000 in Odyssey Blocking Buffer with 0.1% Tween and 0.01% SDS. After three washes in TBST, the membranes were incubated in TBS before imaging in the LICOR Odyssey Fc.

Mass spectrometry and analysis.

Proteins were precipitated with 23% TCA and washed with acetone. Protein pellets solubilized in 8 M urea, 100 mM Tris pH 8.5, and reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride (Sigma-Aldrich, St. Louis, MO, product C4706) and alkylated with 55 mM 2-Chloroacetamide (Fluka Analytical, product 22790). Proteins were digested for 18 h at 37°C in 2 M urea 100 mM Tris pH 8.5, 1 mM $CaCl_2$ with 2 µg trypsin (Promega, Madison, WI, product V5111). Single phase analysis (in replicate) was performed using a Dionex 3000 pump and a Thermo LTQ Orbitrap Velos using an in-house built electrospray stage (Wolters et al., 2001). Protein and peptide identification and protein quantitation were done with Integrated Proteomics Pipeline, IP2 (Integrated Proteomics Applications, Inc., San Diego, CA. http://www.integratedproteomics.com/). Tandem mass spectra were extracted from raw files using RawConverter (He et al., 2015) with monoisotopic peak option and were searched against protein database release WS260 from WormBase, with FLAG-tagged SNPC-4, common contaminants and reversed sequences added, using ProLuCID (Peng et al., 2003; Xu et al., 2006). The search space included all fully-tryptic and half-tryptic peptide candidates with a fixed modification of 57.02146 on C. Peptide candidates were filtered using DTASelect (Tabb et al., 2002).

Using custom R scripts, average enrichment between SNPC-4::3xFlag and no-tag control immunoprecipitation experiments were calculated. For each experiment, enrichment was normalized by dividing the peptide count for each protein by the total peptide count. Adjusted *p*-values were calculated by applying

the Bonferroni method using DESeq2 (Love et al., 2014). Although SNPC-3.1 and SNPC-3.2 are reported to have the same amino acid sequence, we have picked up differential peptide coverage in the *fem-1(-)* mutant for these two proteins and represented them as two different data points.

Immunofluorescence microscopy.

Adult gonads were dissected into egg buffer (25 mM HEPES pH 7.4, 118 mM NaCl, 48 mM KCl, 2 mM EDTA, 0.5 mM EGTA) with 30 mM sodium azide and 0.1% Tween-20, and fixed for 10 s in 1% formaldehyde in egg buffer followed by 1 min in 100% methanol at -20°C. All washing and staining was completed in suspension. Germlines were blocked in normal goat serum or 1x Roche blocking buffer in PBST (PBS + 0.2% Tween) for 30 min at room temperature. Primary mouse anti-Flag (Sigma F1804) and rat anti-Ollas (NBP1-06713SS) antibodies were used at 1:200 in blocking agent in PBST. AlexaFluor 555 goat anti-mouse and AlexaFluor 488 goat anti-rat secondary antibodies (ThermoFisher) were used at 1:400 in blocking agent in PBST. Germlines were stained with 0.5 µg/mL DAPI and then mounted in Vectashield with DAPI (Vector Laboratories H-1200). Images were acquired at 63x on a Zeiss LSM700 confocal microscope. Publication images were acquired at 100x on a GE DeltaVision microscope. Image processing was performed using SoftWoRx to collect 3D image stacks, deconvolve (enhanced ratio, 20 cycles), and compile into a maximum intensity projection. Composite images were stitched and colored in Fiji using the Stitching plugin (Preibisch et al., 2009).

Chromatin immunoprecipitation, library prep, and sequencing.

Worms were grown in liquid culture as previously described (Zanin et al., 2011). 250 µM auxin was added to *snpc-1.3::3xflag; snpc-4::aid::ollas; Psun-1::TIR1* worms 4 h before collection at 48 h post L1 at 20°C. After washing, the gut was cleared for 15 min by nutation in M9, followed by three washes in M9. Worms were live-crosslinked in 2.6% formaldehyde in water for 30 min at room temperature with nutation. Crosslinking was quenched with a final concentration of 125 mM glycine for 5 min with nutation. After three washes with water, worms were flash-frozen in liquid nitrogen. Frozen worm pellets were ground into powder using the Retsch MM40 ball mill homogenizer for 2 rounds of 1 min at 30 Hz. Frozen worm powder was resuspended in 1x RIPA buffer (1x PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) for 10 min at 4°C. Crosslinked chromatin was sonicated using a Diagenode Bioruptor Pico for three 3-min cycles, 30 s on/off. 10 µg chromatin was nutated overnight at 4°C with 2 µg of Flag antibody (Sigma-Aldrich, F1804) and then for 1.5 h with 50 µL mouse IgG Dynabeads (Invitrogen). Input amount was 10% of IP. Chromatin was de-crosslinked and extracted as described previously (Weiser et al., 2017). Individual input and IP samples of each genotype were processed for both sequencing and quantitative PCR.

Libraries were prepared and multiplexed using the Ovation Ultralow Library Systems v2 (NuGEN Technologies) according to the manufacturer's protocol. The Illumina HiSeq 4000 platform was used to generate 50 bp single-end reads for SNPC-1.3 ChIP-seq libraries. The NovaSeq 6000 platform was used to generate 50 bp paired-end reads for TRA-1 ChIP-seq libraries.

Quantitative PCR of ChIP samples.

ChIP DNA was eluted in 18 µL of 1x TE pH 8.0 and 2 µL of 20 mg/mL RNase A (Invitrogen, Thermo Fisher Scientific). For a final reaction volume of 25 µL, each reaction consisted of final 1x Absolute Blue SYBR Green (Thermo Fisher Scientific), 35 nM each of forward and reverse primer, and 2 uL ChIP eluate. Reactions were performed in technical duplicates in a BioRad CF96 Real Time PCR thermal cycler.

Hermaphrodite fertility assays.

Gravid worms (previously maintained at 20°C) were subjected to hypochlorite treatment and their progeny were plated onto NGM at 25°C (P0). At the L2 or L3 stage, worms were singled onto individual plates and their progeny (F1) counted.

Mating assays.

To test male-dependent rescue of *fem-1(hc17)* fertility, 10–12 hermaphrodites of each strain were grown at 20°C and embryos were isolated by allowing egg lay for 2 h before removal. Embryos were shifted to 25°C and upon reaching the L4 stage (24 h), ten *him-8(e1489)* L4 males were transferred and mated with two *fem-1(hc17)* females. Brood size was quantified by counting when a majority of progeny had at least reached the young adult stage (about 3 days after transfer). To test the fertility of the hermaphrodites upon mating, 10–12 hermaphrodites of each strain were grown at 20°C and embryos were isolated after egg lay for 2 h before removal. Embryos were shifted to 25°C and ten *col-19(GFP+)* L4-staged males (24 h) were then transferred with a single hermaphrodite (36 h) and the number of live cross progeny were counted after

reaching adulthood. Brood size was quantified by counting when the majority of progeny had at least reached the young adult stage (about 3 days after transfer).

Sperm activation assay and imaging.

To perform sperm activation assays, spermatids were dissected from adult males that were shifted to 25°C during the embryo stage, and isolated prior to sexual maturity (about 48 h post L1). Dissection was performed directly on glass slides in sperm medium (50 mM HEPES pH 7.8, 50 mM NaCl, 25 mM KCl, 5 mM $CaCl_2$, and 1 mM $MgSO_4$) supplemented with 20 µg/mL pronase E (Millipore Sigma). For the characterization of sperm morphology, sperm were imaged 30 min after the addition of pronase E. Individual sperm were manually categorized into two types: spermatids with normal pseudopods or spermatids with irregular or no pseudopods (Shakes and Ward, 1989). For Figure 2.21, Z stacks were imaged in 10 s intervals for 30 min and a representative in-focus stack was chosen at every 3 min interval. To characterize sperm activation dynamics, sperm were individually followed across 10 s intervals for 30 min and the different stages of sperm activation were designated into four categories based on these morphological changes: 1) undifferentiated spermatid, 2) spiky intermediate characterized by the presence of spike growth, 3) growing or motile pseudopod by the presence of a pseudopod, and 4) immobile sperm when little movement was observed either in the sperm body or pseudopod for longer than 30 s. Statistical significance was assessed using Student's t-test.

Quantitative and statistical analysis

Unless otherwise stated, all quantitative analyses are shown as mean with standard deviation represented as error bars. For qRT-PCR, fertility and mating assays, and western blot, at least 2 independent experiments were performed; one representative biological replicate is shown.

Small RNA-seq analysis.

Raw small RNA-seq reads were trimmed for Illumina adapters and quality (SLIDING WINDOW: 4:25) using Trimmomatic 0.39 (Bolger et al., 2014). Trimmed reads were then filtered using bbmap 38.23 (http://jgi.doe.gov/data-and-tools/bb-tools) to retain reads that were 15–30 nt in length. These filtered reads were aligned to the *C. elegans* WBcel235 (Cunningham et al., 2018) reference genome using Bowtie 1.1.1 (Langmead et al., 2009) with parameters -v 0 -k 5 –best –strata –tryhard. Quality control of raw and aligned reads was performed using FastQC 0.11.7 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), SAMtools 1.9 (Li et al., 2009) and in-house Python and R scripts. Mapped reads were assigned to genomic features using featureCounts from Subread 1.6.3 (Liao et al., 2014), taking into account overlapping and multi-mapping reads (-O -M). Raw counts were normalized within DESeq2 1.26.0 (Love et al., 2014) and principal component analysis (PCA) was performed using the regularized log transform of normalized counts within DESeq2. In addition, we distributed mapped reads by size and 5’ nucleotide identity to verify the presence of small RNA species such as 22G-RNAs and 21U-RNAs.

To identify differentially expressed genes, DESeq2 was applied to piRNAs on chromosome IV. In this study (method 1), we define significant and differentially

expressed genes as having an absolute value of $\log_2$(fold-change) ≥ 0.26 and FDR of ≤ 0.05 (Benjamini-Hochberg). The $\log_2$(fold-change) threshold and significance level were selected based on benchmarking the differential expression results against the Taqman piRNA expression assays. At the chosen cutoffs, differential expression analysis captures the Taqman assays results for the three male piRNAs (21UR-1258, 21UR-3142, and 21UR-3766) and 3 female piRNAs (21UR-1848, 21UR-2502, and 21UR-4817). Contrasts between mutant and wildtype were designed without independent filtering.

For motif discovery, nucleotide sequences were extracted from the reference genome with 60 nt upstream of each piRNA and submitted to the MEME suite 5.1.1 (Bailey et al., 2009). Results from MEME were used to generate the sequence logo plot with the median position of the C-nucleotide of the identified motif, number of piRNAs, and the associated E-value.

A second, independent small RNA-seq analysis workflow was implemented to validate our results. 16–30 nt small RNA sequences were parsed from adapters and reads with >3 nt falling below a quality score of Q30 were discarded. Reads were mapped to the *C. elegans* WS230 (Stein et al., 2001) reference genome using CASHX v. 2.3 (Fahlgren et al., 2009) allowing for 0 mismatches. Custom Perl, Awk, and R scripts were used to count features and to generate PCA and size distribution plots. Multi-mapping reads were assigned proportionally to each possible locus. Differential expression analysis was done using DESeq2 v. 1.18.1 (Love et al., 2014). A reporting threshold was set at an absolute value of $\log_2$(fold-change) ≥ 0.26 and a Benjamini-Hochberg-corrected p ≤ 0.20.

ChIP-seq analysis.

De-multiplexed raw ChIP-seq data in FASTQ format were trimmed for adapters and sequencing quality score > Q25 using Trim Galore! 0.5.0 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and aligned to *C. elegans* reference genome WBcel235 (Cunningham et al., 2018) using Bowtie2 2.3.4.2 (Langmead and Salzberg, 2012) with default parameters. Post-alignment filtering was then performed to remove PCR duplicates using the MarkDuplicates utility within Picard 2.22.1 (http://broadinstitute.github.io/picard/). In addition, SAMtools 1.9 was applied to remove unmapped reads and reads that mapped with MAPQ 30 but were not of primary alignment or failed sequence platform quality checks (SAMtools -F 1804 -q 30) (Li et al., 2009).

To identify and visualize binding sites and peaks for SNPC-1.3 ChIP-seq, filtered SNPC-1.3 ChIP-seq reads were extended to 200 bp to account for the average length of ChIP fragments. We then partitioned the genome into consecutive, non-overlapping 1 kb bins and calculated read coverage, normalized by sequencing depth of each library, based on the total read count in each bin. Bins with read coverages in the IP sample that fell below the median read coverage of piRNA-depleted bins on chromosome IV in the relevant input control were excluded from further analysis. Bins containing only male, female, and non-enriched piRNAs (as defined by small RNA-seq analysis) were then extracted to generate binding profiles and heatmaps. For this, the bamCompare tool in deepTools 3.3.1 (Ramírez et al., 2016) was used to calculate the ratio between read coverage of each ChIP sample and input control (--scaleFactorsMethod None

--normalizeUsing CPM --operation ratio --binSize 50 --ignoreForNormalization MtDNA --extendReads 200). The ENCODE ce11 blacklist was also supplied (https://github.com/Boyle-Lab/Blacklist/tree/master/lists). The bamCompare output was then used in deepTools computeMatrix to calculate scores for plotting profiles and heatmaps with deepTools plotProfile and plotHeatmap.

TRA-1 ChIP-seq peaks were called by callpeak within MACS 2.1.2 (Zhang et al., 2008) (--pvalue 0.05) with filtered TRA-1 ChIP-seq reads and relevant input controls. TRA-1 signal tracks were generated by calculating fold enrichment from read count-normalized genome-wide pileup and lambda track outputs by callpeak (bdgcmp in MACS2). The ENCODE ce 11 blacklist was supplied in this analysis (https://github.com/Boyle-Lab/Blacklist/tree/master/lists). The bamCompare tool in deepTools 3.3.1 (Ramírez et al., 2016) was used to quantify read coverage of each ChIP sample and input control.

Reproducibility between SNPC-1.3 and TRA-1 ChIP-seq replicates was assessed by applying deepTools bamCompare, as described above, and deepTools plotCorrelation to depict pairwise correlations between replicates and compute the Pearson correlation coefficient.

2.9 Acknowledgements

We thank Himani Galagali and Natasha Weiser for helpful comments on the manuscript. We thank members of the Kim Lab (Amelia Alessi, Mindy Clark, Gregory Fuller, Jessica Kirshner, Alex Rittenhouse, Darius Mostaghimi, Lars Benner), Tatjana Trcek, Jocelyn Haversat, Yumi Kim, Angela Andersen, Aurelia Mapps, and Jacqueline Tay for helpful suggestions. Computational resources

were provided by the Maryland Advanced Research Computing Center (MARCC). Some strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH Office of Research Infra-structure Programs (P40 OD010440). This work was supported by grants from the NSF DGE-1746891 (to R.J.T.), NIH R35 GM130272 (to S.E.J); NIH R35 GM119775 (to T.A.M.); and NIH R01 GM129301 and NIH R01 GM118875 (to J.K.K.).

2.10 References

Andersen, P.R., Tirian, L., Vunjak, M., and Brennecke, J. (2017). A heterochromatin-dependent transcription machinery drives piRNA expression. Nature *549*, 54-59.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during *Drosophila melanogaster* development. Dev. Cell *5*, 337-350.

Armisen, J., Gilchrist, M.J., Wilczynska, A., Standart, N., and Miska, E.A. (2009). Abundant and dynamically expressed miRNAs, piRNAs, and other small RNAs in the vertebrate *Xenopus tropicalis*. Genome Res. *19*, 1766-1775.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J.L., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. Cell *150*, 88-99.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: Tools for motif discovery and searching. Nucleic Acids Res. *37*, W202-8.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. Mol. Cell *31*, 67-78.

Beanan, M.J., and Strome, S. (1992). Characterization of a germ-line proliferation mutation in *C. elegans*. Development *116*, 755-766.

Beltran, T., Barroso, C., Birkle, T.Y., Stevens, L., Schwartz, H.T., Sternberg, P.W., Fradin, H., Gunsalus, K., Piano, F., Sharma, G., et al. (2019). Comparative epigenomics reveals that RNA polymerase II pausing and chromatin domain organization control nematode piRNA biogenesis. Dev. Cell *48*, 793-810.

Berkseth, M., Ikegami, K., Arur, S., Lieb, J.D., and Zarkower, D. (2013). TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination. Proc. Natl. Acad. Sci. USA *110*, 16033-16038.

Bessereau, J.-L. (2006). Transposons in *C. elegans*. In Wormbook, The *C. elegans* Research Community, WormBook ed., doi/10.1895/wormbook.1.70.1, http://www.wormbook.org.

Billi, A.C., Freeberg, M.A., Day, A.M., Chun, S.Y., Khivansara, V., and Kim, J.K. (2013). A conserved upstream motif orchestrates autonomous, germline-enriched expression of *Caenorhabditis elegans* piRNAs. PLoS Genet. *9*, e1003392.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. Cell *128*, 1089-1103.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science *322*, 1387-1392.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. Genetics *77*, 71-94.

Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E.J., Roloff, T.C., Beisel, C., Schübeler, D., Stadler, M.B., and Peters, A.H.F.M. (2010). Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat. Struct. Mol. Biol. *17*, 679-687.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes

multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447-451.

Carmell, M.A., Girard, A., van de Kant, H.J.G., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell *12*, 503-514.

Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012). Promoters recognized by Forkhead proteins exist for individual 21U-RNAs. Mol. Cell *47*, 734-745.

Chen, P., and Ellis, R.E. (2000). TRA-1 regulates transcription of *fog-3*, which controls germ cell fate in *C. elegans*. Development *127*, 3119-3129.

Chen, Y.-C.A., Stuwe, E., Luo, Y., Ninova, M., Le Thomas, A., Rozhavskaya, E., Li, S., Vempati, S., Laver, J.D., Patel, D.J., et al. (2016). Cutoff suppresses RNA polymerase II termination to ensure expression of piRNA precursors. Mol. Cell *63*, 97-109.

Clarke, N.D., and Berg, J.M. (1998). Zinc fingers in *Caenorhabditis elegans*: Finding families and probing pathways. Science *282*, 2018-2022.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by *piwi* are essential for stem cell self-renewal. Genes Dev. *12*, 3715-3727.

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R., Bhai, J., Billis, K., Boddu, S., et al. (2018). Ensembl 2019. Nucleic Acids Res. *47*, D745-D751.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. Mol. Cell *31*, 79-90.

Deng, W., and Lin, H. (2002). *miwi*, a murine homolog of *piwi*, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell *2*, 819-830.

Di Giacomo, M., Comazzetto, S., Saini, H., De Fazio, S., Carrieri, C., Morgan, M., Vasiliauskaite, L., Benes, V., Enright, A.J., and O'Carroll, D. (2013). Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell *50*, 601-608.

Doniach, T., and Hodgkin, J. (1984). A sex-determining gene, *fem-1*, required for both male and hermaphrodite development in *Caenorhabditis elegans*. Dev. Biol. *106*, 223-235.

Fahlgren, N., Sullivan, C.M., Kasschau, K.D., Chapman, E.J., Cumbie, J.S., Montgomery, T.A., Gilbert, S.D., Dasenko, M., Backman, T.W.H., Givan, S.A., et al. (2009). Computational and analytical framework for small RNA profiling by high-throughput sequencing. RNA *15*, 992-1002.

Gainetdinov, I., Colpan, C., Arif, A., Cecchini, K., and Zamore, P.D. (2018). A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. Mol. Cell *71*, 775-790.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Goh, W.-S.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. Genes Dev. *28*, 797-807.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709-1714.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D. Jr, and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. Cell *151*, 1488-1500.

Hammoud, S.S., Low, D.H.P., Yi, C., Carrell, D.T., Guccione, E., and Cairns, B.R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. Cell Stem Cell *15*, 239-253.

Harris, A.N., and Macdonald, P.M. (2001). *aubergine* encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. Development *128*, 2823-2832.

He, L., Diedrich, J., Chu, Y.-Y., and Yates, J.R. III. (2015). Extracting accurate precursor information for tandem mass spectra by RawConverter. Anal. Chem. *87*, 11361-11367.

Henry, R.W., Sadowski, C.L., Kobayashi, R., and Hernandez, N. (1995). A TBP-TAF complex required for transcription of human snRNA genes by RNA polymerases II and III. Nature *374*, 653-656.

Henry, R.W., Mittal, V., Ma, B., Kobayashi, R., and Hernandez, N. (1998). SNAP19 mediates the assembly of a functional core promoter complex (SNAPc) shared by RNA polymerases II and III. Genes Dev. *12*, 2664-2672.

Hodgkin, J. (1987). A genetic analysis of the sex-determining gene, *tra-1*, in the nematode *Caenorhabditis elegans*. Genes Dev. *1*, 731-745.

Hodgkin, J., Horvitz, H.R., and Brenner, S. (1979). Nondisjunction mutants of the nematode *Caenorhabditis elega*ns. Genetics *91*, 67-94.

Hung, K.-H., and Stumph, W.E. (2010). Regulation of snRNA gene expression by the *Drosophila melanogaster* small nuclear RNA activating protein complex (DmSNAPc). Crit. Rev. Biochem. Mol. Biol. *46*, 11-26.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Jawdekar, G.W., and Henry, R.W. (2008). Transcriptional regulation of human small nuclear RNA genes. Biochim. Biophys. Acta *1779*, 295-305.

Kamath, R.S., and Ahringer, J. (2003). Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods *30*, 313-321.

Kaneshiro, K.R., Rechtsteiner, A., and Strome, S. (2019). Sperm-inherited H3K27me3 impacts offspring transcription and development in *C. elegans*. Nat. Commun. *10*, 1271.

Kasper, D.M., Wang, G., Gardner, K.E., Johnstone, T.G., and Reinke, V. (2014). The *C. elegans* SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell *31*, 145-158.

Kato, M., de Lencastre, A., Pincus, Z., and Slack, F.J. (2009). Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. Genome Biol. *10*, R54.

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J.S., Zhang, F., Schultz, N., Koppetsch, B.S., Nowosielska, A., et al. (2009). The *Drosophila* HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell *138*, 1137-1149.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. *22*, 908-917.

Lamont, L.B., and Kimble, J. (2007). Developmental expression of FOG-1/CPEB protein and its control in the *C. elegans* hermaphrodite germ line. Dev. Dyn. *236*, 871-879.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357-359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Levine, M., Cattoglio, C., and Tjian, R. (2014). Looping back to leap forward: transcription enters a new era. Cell *157*, 13-25.

Li, C., Harding, G.A., Parise, J., McNamara-Schroeder, K.J., and Stumph, W.E. (2004). Architectural arrangement of cloned proximal sequence element-binding protein subunits on *Drosophila* U1 and U6 snRNA gene promoters. Mol. Cell. Biol. *24*, 1897-1906.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z., et al. (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol. Cell *50*, 67-81.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. Development *124*, 2463-2476.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Ma, B., and Hernandez, N. (2002). Redundant cooperative interactions for assembly of a human U6 transcription initiation complex. Mol. Cell Biol. *22*, 8067-8078.

Mittal, V., Ma, B., and Hernandez, N. (1999). SNAPc: a core promoter factor with a built-in DNA-binding damper that is deactivated by the Oct-1 POU domain. Genes Dev. *13*, 1807-1821.

Moissiard, G., Bischof, S., Husmann, D., Pastor, W.A., Hale, C.J., Yen, L., Stroud, H., Papikian, A., Vashisht, A.A., Wohlschlegel, J.A., et al. (2014). Transcriptional gene silencing by *Arabidopsis microrchidia* homologues involves the formation of heteromers. Proc. Natl. Acad. Sci. USA *111*, 7474-7479.

Mohn, F., Sienski, G., Handler, D., and Brennecke, J. (2014). The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand piRNA clusters in *Drosophila*. Cell *157*, 1364-1379.

Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. Genes Dev. *21*, 682-693.

Ozata, D.M., Gainetdinov, I., Zoch, A., O'Carroll, D., and Zamore, P.D. (2019). PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. *20*, 89-108.

Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High efficiency, homology-directed genome editing in *Caenorhabditis elegans* using CRISPR-Cas9 ribonucleoprotein complexes. Genetics *201*, 47-54.

Pane, A., Jiang, P., Zhao, D.Y., Singh, M., and Schüpbach, T. (2011). The Cutoff protein regulates piRNA cluster expression and piRNA production in the *Drosophila* germline. EMBO J. *30*, 4601-4615.

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC−MS/MS) for large-scale protein analysis: the yeast proteome. J. Proteome Res. *2*, 43-50.

Pires-daSilva, A., and Sommer R.J., (2004). Conservation of the global sex determination gene *tra-1* in distantly related nematodes. Genes Dev. *18*, 1198-1208.

Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. Bioinformatics *25*, 1463–1465.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160-W165.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell *127*, 1193-1207.

Shakes, D.C., and Ward, S. (1989). Initiation of spermiogenesis in *C. elegans*: A pharmacological and genetic analysis. Dev. Biol. *134*, 189-200.

Shen, E.-Z., Chen, H., Ozturk, A.R., Tu, S., Shirayama, M., Tang, W., Ding, Y.-H., Dai, S.-Y., Weng, Z., and Mello, C.C. (2018). Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the *C. elegans* germline. Cell *172*, 937-951.

Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D. Jr, and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. Cell *150*, 65-77.

Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., and Spieth, J. (2001). WormBase: network access to the genome and biology of *Caenorhabditis elegans*. Nucleic Acids Res. *29*, 82-86.

Su, Y., Song, Y., Wang, Y., Jessop, L., Zhan, L., and Stumph, W.E. (1997). Characterization of a *Drosophila* proximal-sequence-element-binding protein involved in transcription of small nuclear RNA genes. Eur. J. Biochem. *248*, 231-237.

Tabb, D.L., McDonald, W.H., and Yates, J.R. III. (2002). DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. *1*, 21-26.

Tabuchi, T.M., Rechtsteiner, A., Jeffers, T.E., Egelhofer, T.A., Murphy, C.T., and Strome, S. (2018). *Caenorhabditis elegans* sperm carry a histone-based epigenetic memory of both spermatogenesis and oogenesis. Nat. Commun. *9*, 4310.

Xu, T., Venable, J.D., Park, S.K., Cociorva, D., Lu, B., Liao, L., Wohlschlegel, J., Hewel, J., Yates, J.R. III. (2006). ProLuCID, a fast and sensitive tandem

mass spectra-based protein identification program. Mol. Cell. Proteom. *5*, S174.

Timmons, L., and Fire, A. (1998). Specific interference by ingested dsRNA. Nature *395*, 854.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr. Biol. *18*, 861-867.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R.A., Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E.A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in *C. elegans*. Genes Dev. *28*, 783-796.

Weiser, N.E., Yang, D.X., Feng, S., Kalinava, N., Brown, K.C., Khanikar, J., Freeberg, M.A., Snyder, M.J., Csankovszki, G., Chan, R.C., et al. (2017). MORC-1 integrates nuclear RNAi and transgenerational chromatin architecture to promote germline immortality. Dev. Cell *41*, 408-423.e7.

Weng, C., Kosalka, J., Berkyurek, A.C., Stempor, P., Feng, X., Mao, H., Zeng, C., Li, W.-J., Yan, Y.-H., Dong, M.-Q., et al. (2019). The USTC co-opts an ancient machinery to drive piRNA transcription in *C. elegans*. Genes Dev. *33*, 90-102.

Williams, Z., Morozov, P., Mihailovic, A., Lin, C., Puvvula, P.K., Juranek, S., Rosenwaks, Z., and Tuschl, T. (2015). Discovery and characterization of piRNAs in the human fetal ovary. Cell Rep. *13*, 854-863.

Wolters, D.A., Washburn, M.P., and Yates, J.R. III. (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal. Chem. *73*, 5683-5690.

Wong, M.W., Henry, R.W., Ma, B., Kobayashi, R., Klages, N., Matthias, P., Strubin, M., and Hernandez, N. (1998). The large subunit of basal transcription factor SNAPc is a Myb domain protein that interacts with Oct-1. Mol. Cell Biol. *18*, 368-377.

Yang, Q., Hua, J., Wang, L., Xu, B., Zhang, H., Ye, N., Zhang, Z., Yu, D., Cooke, H.J., Zhang, Y., et al. (2013). MicroRNA and piRNA profiles in normal human testis detected by next generation sequencing. PLoS ONE *8*, e66809.

Yoon, J.B., Murphy, S., Bai, L., Wang, Z., and Roeder, R.G. (1995). Proximal sequence element-binding transcription factor (PTF) is a multisubunit complex required for transcription of both RNA polymerase II- and RNA polymerase III-dependent small nuclear RNA genes. Mol. Cell Biol. *15*, 2019-2027.

Zanin, E., Dumont, J., Gassmann, R., Cheeseman, I., Maddox, P., Bahmanyar, S., Carvalho, A., Niessen, S., Yates, J.R. III., Oegema, K., et al. (2011). Affinity purification of protein complexes in *C. elegans*. Methods Cell Biol. *106*, 289-322.

Zarkower, D., and Hodgkin, J. (1993). Zinc fingers in sex determination: only one of the two *C. elegans* Tra-1 proteins binds DNA *in vitro*. Nucleic Acids Res. *21*, 3691-3698.

Zhang, D., Tu, S., Stubna, M., Wu, W.-S., Huang, W.-C., Weng, Z., and Lee, H.-C. (2018). The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science *359*, 587-592.

Zhang, L., Ward, J.D., Cheng, Z., and Dernburg, A.F. (2015a). The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in *C. elegans*. Development *142*, 4374-4384.

Zhang, P., Kang, J.-Y., Gou, L.-T., Wang, J., Xue, Y., Skogerboe, G., Dai, P., Huang, D.-W., Chen, R., Fu, X.-D., et al. (2015b). MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. Cell Res. *25*, 193-207.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhou, X., Zuo, Z., Zhou, F., Zhao, W., Sakaguchi, Y., Suzuki, T., Suzuki, T., Cheng, H., and Zhou, R. (2010). Profiling sex-specific piRNAs in zebrafish. Genetics *186*, 1175-1185.

Chapter 3

A kinesin Klp10A mediates cell cycle-dependent shuttling of Piwi between nucleus and nuage.

3.1 Citation

Zsolt G. Venkei1, Charlotte Choi4, Suhua Feng5,6, Cuie Chen1, 3, Steven E. Jacobsen 5,6,7, John K. Kim4, and Yukiko M. Yamashita1, 2, 3, * (2020). A kinesin Klp10A mediates cell cycle-dependent shuttling of Piwi between nucleus and nuage. PLOS Genetics 16(10):e1009147.

3.2 Author Contributions

Conceptualization:	Zsolt G. Venkei, John K. Kim, Yukiko M. Yamashita.
Data curation:	Charlotte Choi, John K. Kim.
Formal analysis:	Zsolt G. Venkei, Charlotte Choi.
Investigation:	Zsolt G. Venkei, Charlotte Choi, Suhua Feng, Cuie Chen.
Methodology:	Zsolt G. Venkei, Charlotte Choi, John K. Kim, Yukiko M. Yamashita.
Writing:	Zsolt G. Venkei, Charlotte Choi, John K. Kim, Yukiko M. Yamashita.

3.3 Abstract

The piRNA pathway protects germline genomes from selfish genetic elements (e.g. transposons) through their transcript cleavage in the cytoplasm

and/or their transcriptional silencing in the nucleus. Here, we describe a mechanism by which the nuclear and cytoplasmic arms of the piRNA pathway are linked. We find that during mitosis of *Drosophila* spermatogonia, nuclear Piwi interacts with nuage, the compartment that mediates the cytoplasmic arm of the piRNA pathway. At the end of mitosis, Piwi leaves nuage to return to the nucleus. Dissociation of Piwi from nuage occurs at the depolymerizing microtubules of the central spindle, mediated by a microtubule-depolymerizing kinesin, Klp10A. Depletion of *klp10A* delays the return of Piwi to the nucleus and affects piRNA production, suggesting the role of nuclear-cytoplasmic communication in piRNA biogenesis. We propose that cell cycle-dependent communication between the nuclear and cytoplasmic arms of the piRNA pathway may play a previously unappreciated role in piRNA regulation.

3.4 Introduction

Piwi-associated RNAs (piRNAs), a class of endogenous small RNAs found across many organisms, associate with Piwi proteins of the Argonaute family to silence active transposable elements (TEs)(Aravin et al., 2003; Batista et al., 2008; Brennecke et al., 2007; Girard et al., 2006; Ruby et al., 2006). In *Drosophila melanogaster*, Piwi, Aubergine (Aub), and Argonaute 3 (Ago3) comprise the three Piwi proteins of the piRNA pathway (Brennecke et al., 2007; Gunawardane et al., 2007; Nishida et al., 2007; Wang et al., 2015). The piRNA pathway is initiated in the nucleus, where primary piRNA precursors are transcribed from genomic piRNA clusters, translocated to the cytoplasm, and processed into mature piRNAs (Brennecke et al., 2007; Vagin et al., 2006). In the cytoplasm of the germline,

primary piRNAs are amplified in a “ping-pong” cycle that simultaneously cleaves TEs and produces secondary piRNAs (Gunawardane et al., 2007; Li et al., 2009; Wang et al., 2014). Specifically, primary piRNAs antisense to TEs are loaded into and direct Aub to cleave TE transcripts, which produces secondary piRNAs that are subsequently loaded into Ago3. The secondary piRNAs generated by the ping-pong cycle are then loaded into Piwi and translocated into the nucleus as a piRNA-induced silencing complex (piRISC) (Senti et al., 2015; Wang et al., 2015). Within the nucleus, Piwi leads to the transcriptional repression of TEs through the deposition of heterochromatic histone marks (Klenov et al., 2014; Le Thomas et al., 2013; Sienski et al., 2015). Thus, the cytoplasmic arm of the piRNA pathway intimately interacts with the nuclear arm to affect silencing at the post-transcriptional and transcriptional levels, respectively.

As evidence of the tight coupling between the two arms of the piRNA pathway, ping-pong amplification takes place in nuage – the electron-dense, phase-separated granules that are anchored directly to the cytoplasmic face of the nuclear pore (Lim and Kai, 2007; Malone et al., 2009; Zhang et al., 2012). The interface between nuage and the nucleus has been shown to be critical for piRNA biogenesis and transposon target recognition and repression. Although nuage has been characterized as a static and long-term platform for piRNA biogenesis, individual nuage components, such as Aub and Ago3, are dynamically shuttled in and out of nuage (Andress et al., 2016; Brennecke et al., 2007; Ryazansky et al., 2016; Snee and Macdonald, 2004; Wang et al., 2015). Unlike Aub and Ago3, Piwi shows predominantly nuclear localization (Gonzalez et al., 2015) and has not been

well-characterized within the context of nuage. Although previous studies have identified multiple piRNA pathway components that are required for proper localization of piRNA amplification machinery to nuage, including Spindle-E, Qin, Krimper, and Vasa, the underlying mechanism by which ping-pong-amplified piRNAs translocate into the nucleus remains poorly understood (Andress et al., 2016; Sato et al., 2015; Wang et al., 2015; Webster et al., 2015; Zhang et al., 2012; Zhang et al., 2011).

Here we show that nuclear and cytoplasmic components of the piRNA pathway interact specifically during mitosis of *Drosophila* spermatogonia (SGs) to likely facilitate communication between the nuclear and cytoplasmic arms of the pathway. We found that Piwi localizes to nuage specifically during mitosis, and returns to the nucleus upon mitotic exit. Through immunoprecipitation and mass spectrometry, we found that Klp10A, a microtubule (MT)-depolymerizing kinesin motor protein (Rogers et al., 2004), physically interacts with components of the piRNA pathway. In the absence of Klp10A, Piwi remains in nuage for a prolonged period after mitotic exit, thus delaying its return to the nucleus. Cytological observations suggest that dissociation of Piwi from nuage at the end of mitosis occurs at the central spindle while it is being depolymerized in a klp10A-dependent manner. *klp10A*-depletion leads to increased piRNA production, supporting the physiological role of Klp10A-mediated regulation of Piwi localization. We propose that the interaction of Piwi with nuage during mitosis may represent a previously unappreciated mechanism to regulate piRNA biogenesis.

3.5 Results

3.5.1 Klp10A physically interacts with piRNA pathway components in germline stem cells and spermatogonia of *Drosophila* testis.

Drosophila spermatogenesis is supported by asymmetrically dividing germline stem cells (GSCs), which produce a self-renewing GSC and a gonialblast (GB) that initiates the differentiation program. GBs further undergo four rounds of mitotic divisions as spermatogonia (SGs), which then enter meiotic program as spermatocytes (SCs) (Fig 3.1A). Asymmetric divisions of *Drosophila* male GSCs, as well as several other systems, are mediated by stereotypical positioning of the centrosomes (Cheng et al., 2008; Conduit and Raff, 2010; Inaba et al., 2015b; Januschke et al., 2011; Januschke et al., 2013; Salzmann et al., 2014; Venkei and Yamashita, 2015; Wang et al., 2009; Yamashita et al., 2003; Yamashita et al., 2007). Therefore, stem cell-specific centrosomal components are of significant interest in understanding the mechanisms that regulate asymmetric stem cell divisions.

In a previous study, we showed that Klp10A protein is enriched on the centrosomes specifically in GSCs, but not in differentiating SGs (Chen et al., 2016) (Fig 3.1A). In an attempt to isolate proteins that specifically localize to GSC centrosomes, we affinity-purified Klp10A from a GSC-enriched extract, using either specific antibody or anti-GFP antibody combined with expression of Klp10A-GFP (see Methods). To enrich for GSCs, the self-renewal factor Unpaired (Upd) was expressed (*nos-gal4>UAS-upd*). Upd is a ligand expressed by hub cells to activate the JAK-STAT pathway, which determines GSC self-renewal. It was shown that

ectopic expression of Upd in germ cells (*nos-gal4>UAS-upd*) is sufficient to cause GSC-like tumors in the testis (Kiger et al., 2001; Tulina and Matunis, 2001). Anti-Klp10A and anti-GFP immunoprecipitates were analyzed by mass-spectrometry (see Methods). Many MT-associated proteins were enriched in Klp10A pulldown samples, validating the specificity of the Klp10A pull down experiments.

To our surprise, we found that components of the piRNA pathway, such as Vasa, Aub and Piwi, were enriched in the Klp10A pull-downs. The fact that many piRNA pathway components were consistently identified in Klp10A-pulldowns led us to speculate this interaction might be of significance. Indeed, immunoprecipitation using anti-Klp10A antibody and GSC-enriched extracts confirmed that Klp10A physically interacts with Vasa (Figure 3.1B). Similarly, anti-GFP antibody pull down of GFP-Aub confirmed the Aub-Klp10A interaction (Figure 3.1C). Although it remains elusive whether Klp10A directly binds to these proteins or perhaps through an intermediary such as RNA, the fact that Klp10A co-immunoprecipitates multiple piRNA pathway components indicates that their interaction may be of biological relevance, regardless of whether the interaction is direct or not. Similar interactions between Klp10A and piRNA pathway components (Vasa, Piwi, Aub, Ago3) were observed when extracts from SG tumor was used (nos>dpp, (Kawase et al., 2004; Schulz et al., 2004; Shivdasani and Ingham, 2003) (Figure 3.1D–E), suggesting that the Klp10A interaction with piRNA pathway components is not unique to GSCs. Cytological data confirmed that Klp10A interaction with piRNA pathway components is not limited to GSC but also occurs in SGs (see below). Accordingly, the study diverged from our initial intention to

isolate GSC centrosome-specific components. However, we investigated the unexpected role of klp10A in the regulation of the piRNA pathway, as the study revealed an unappreciated mode of regulation of piRNA biogenesis.

3.5.2 Depletion of klp10A results in alteration of piRNA biogenesis.

To explore whether Klp10A interaction with piRNA pathway components was functionally significant, I examined whether klp10A was required for piRNA biogenesis (Toth et al., 2016). To characterize the role of klp10A in the piRNA pathway, we first performed deep sequencing of small RNAs from wild type testes or testes with germline-specific *klp10A*-knockdown (*nos-gal4>UAS-klp10ATRiP.HMS00920*, validated in our previous study (Chen et al., 2016), and hereafter referred to as klp10ARNAi). I defined piRNAs as reads of 23-29 nucleotides (nt) in length that did not map to microRNAs or ribosomal RNAs. The majority of reads that are 23-29 nucleotides in length start with a Uridine at the 5' most position, which has shown to be a unique, conserved feature of mature piRNAs, whereas reads that are 20-22 nucleotides in length, defined as siRNAs in our study, do not show such enrichment, validating our analysis.

To assess the global changes in piRNA expression upon klp10A loss, I profiled piRNA reads across the *Drosophila* transcriptome. *klp10A*^{RNAi} caused specific upregulation of piRNAs mapping to repetitive elements and piRNA clusters, with minimal differences in piRNAs mapping to other genomic classes, or in other types of small RNAs such as microRNAs (Figure 3.2A). These data suggest that *klp10A* may play a specific role in the piRNA pathway. When I analyzed the abundance of piRNAs mapping to TEs upon loss of *klp10A*, I

observed significant upregulation of piRNA reads both antisense and sense to TEs (Figure 3.2B-E) without any noticeable changes in their distribution (Figure 3.2F). The total amount of piRNA per TE was upregulated up to ~4 times (Figure 3.2B), with upregulation of individual piRNA being more prominent (Figure 3.2C).

The above data showed that both primary and secondary piRNAs were upregulated, suggesting that the ping-pong cycle itself remains intact in *klp10A*^{RNAi}. To test this, I calculated the number of piRNAs with a ping-pong signature (i.e. antisense and sense read pairs that have 10-nucleotide complementary base pairing from their 5' ends) in both wildtype and *klp10A*^{RNAi} libraries. We observed a significant bias for piRNA read pairs with 10-nucleotide complementarity in both wildtype and *klp10A*^{RNAi}, suggesting that the ping-pong pathway indeed remains active upon loss of klp10A. Moreover, we found no significant changes in ping-pong ratios in klp10ARNAi compared to wildtype, further confirming an intact ping-pong cycle.

To address if piRNA upregulation in *klp10A*^{RNAi} leads to changes in TE expression, we performed mRNA-sequencing in wildtype and *klp10A*^{RNAi} germ cells. We found that most TEs did not show any significant changes in expression levels except for a small subset of TEs, which showed moderate downregulation upon loss of *klp10A* (Figure 3.2D). There were no clear common characteristics among these TEs that may explain as to why these TEs exhibits downregulation. Taken together, our results show that klp10A is involved in piRNA production and that Klp10A interaction with piRNA pathway components may have functional significance (see Discussion), prompting us to further examine the underlying

mechanisms. It should be noted that *klp10A*^{RNAi} does not lead to noticeable changes in overall cellular composition of the testes, the process of differentiation (Chen et al., 2016), or cell cycle distribution (see below). Therefore, it is unlikely that skewed composition of cell types/cell cycle is the cause of observed changes in piRNA levels in *klp10A*^{RNAi}.

3.5.3 Klp10A colocalizes with piRNA pathway components at the central spindle during telophase in GSCs and SGs.

To begin to explore how Klp10A might contribute to piRNA biogenesis, we examined the localization of Klp10A protein and piRNA pathway components throughout the cell cycle of GSCs and SGs. Vasa and Aub showed well-established perinuclear localization in nuage throughout interphase (Figure 3.3A, B, E, and F) (Kibanov et al., 2011; Nagao et al., 2010; Nishida et al., 2007). During this period, Klp10A showed centrosome localization in GSCs (Fig 3.3A, E and I), or uniform cytoplasmic localization in SGs (Figure 3.3B, F and J) as we reported previously (Chen et al., 2016). We did not observe clear co-localization between centrosomal Klp10A and nuage, suggesting that the potential role of Klp10A in piRNA pathway is not related to its GSC-specific centrosomal localization.

When cells enter mitosis, nuage became somewhat larger in size, which we refer to as ‘mitotic nuage’ hereafter (Figure 3.3C, G and K). Mitotic nuage appears to be the same population described previously as ‘Vasa granule near the mitotic chromosomes’ by Pek and Kai (Pek and Kai, 2011). At this point, Klp10A was observed at the spindle pole (and weakly on the spindle) as previously reported

(Chen et al., 2016), showing little colocalization with Vasa/Aub (Figure 3.3C, G, K and M–O).

Although Klp10A did not noticeably colocalize with any of nuage components until anaphase, colocalization became clear during telophase: nuage was associated with Klp10A near the center of the dividing cell (Figure 3.3D, H and L–O), which we confirmed to be bundles of central spindle MTs. Importantly, Klp10A localization to the central spindle was observed both in GSCs and SGs (Chen et al., 2016), and colocalization of Klp10A and piRNA pathway components was commonly observed both in GSCs and SGs.

In contrast to Aub and Ago3, which reside in nuage together with Vasa to function in the cytoplasmic arm of the piRNA pathway, Piwi functions in the nucleus to repress TEs at the transcriptional level (Aravin et al., 2008; Le Thomas et al., 2013; Nishida et al., 2015; Toth et al., 2016; Xiol et al., 2014; Zhang et al., 2012). We confirmed that GFP-Piwi localizes to the nucleus during interphase of GSCs and SGs as described previously (Cox et al., 2000; Gonzalez et al., 2015) (Figure 3.3I and J). Interestingly, we found that once cells enter mitosis, GFP-Piwi localized to mitotic nuage together with the components of the cytoplasmic arm of the piRNA pathway, such as Aub and Vasa (Figure 3.3K). During telophase, GFP-Piwi still colocalized with nuage components at the central spindle (Figure 3.3 L, O), after which it returned to the nucleus. Taken together, these data show that piRNA pathway components likely interact with Klp10A at the central spindle during telophase.

3.5.4 Klp10A is required for relocation of Piwi from mitotic nuage to the nucleus at the end of mitosis.

Based on the above results that suggest Klp10A interaction with nuage components at the central spindle, we explored how this interaction may impact the piRNA pathway. To this end, we first examined the localization of Piwi and Vasa in control vs. klp10ARNAi GSCs/SGs.

We confirmed the above observation that Piwi localizes to mitotic nuage by examining the colocalization of GFP-Piwi with Vasa in wild type cells. Although their localization was distinct during interphase (i.e. Vasa in nuage, Piwi in the nucleus) (Figure 3.3, Figure 3.4A and C), Vasa and Piwi colocalized during mitosis (Figure 3.4B), and Piwi returned to the nucleus at the end of mitosis (Figure 3.4C). We further detected physical interaction and colocalization between Vasa and Piwi specifically when cells were blocked in metaphase with MG132 (a proteasome inhibitor that arrests cells in mitosis (Moutinho-Pereira et al., 2010)), or with colcemid (Venkei and Yamashita, 2015). These results suggest that Piwi gains access to nuage specifically during mitosis.

Next, we investigated whether *klp10A*^{RNAi} testes show any defects in the behavior of nuage or Piwi during the cell cycle. Nuage morphology and composition appeared unperturbed in most interphase cells (Figure 3.4E), and in all mitotic cells until anaphase in both GSCs and SGs of *klp10A*^{RNAi} testes (Figure 3.4F). However, the difference between the control and klp10ARNAi germ cells became clear upon exiting mitosis (‘G1’ in Figure 3.4C, 3.3G). In control germ cells, Piwi returned to the nucleus at the end of telophase and became exclusively

nuclear in the subsequent interphase as described above (Figure 3.4C). In contrast, in *klp10A*^{RNAi} germ cells, Piwi persisted in cytoplasmic nuage even after the completion of mitosis (Figure 3.4G). Piwi-Vasa interaction was detectable by co-immunoprecipitation in klp10ARNAi without being arrested in mitosis (Figure 3.4H), consistent with persistent colocalization of Piwi with Vasa at nuage in early G1/S cells (Figure 3.4G). Likely reflecting the delayed/incomplete translocation of Piwi from mitotic nuage to the nucleus, we also observed that the amount of nuclear Piwi was slightly but significantly reduced in GSCs/SGs in *klp10A*^{RNAi} germ cells. Delayed dissociation of Piwi from mitotic nuage at the end of mitosis was confirmed by live observation of GFP-Piwi and mCherry-Vasa expressed in mitotic SGs. Taken together, these results show that 1) Piwi interacts with nuage specifically in mitosis, and 2) klp10A is required for releasing Piwi from mitotic nuage to allow its return to the nucleus at the end of mitosis.

Next, we investigated whether *klp10A*^{RNAi} testes show any defects in the behavior of nuage or Piwi during the cell cycle. Nuage morphology and composition appeared unperturbed in most interphase cells (Figure 3.4E), and in all mitotic cells until anaphase in both GSCs and SGs of *klp10A*^{RNAi} testes (Figure 3.4F). However, the difference between the control and *klp10A*^{RNAi} germ cells became clear upon exiting mitosis (‘G1’ in Figure 3.4C, G). In control germ cells, Piwi returned to the nucleus at the end of telophase and became exclusively nuclear in the subsequent interphase as described above (Figure 3.4C). In contrast, in *klp10A*^{RNAi} germ cells, Piwi persisted in cytoplasmic nuage even after the completion of mitosis (Figure 3.4G). Piwi-Vasa interaction was detectable by

co-immunoprecipitation in *klp10A*^{RNAi} without being arrested in mitosis (Figure 3.4H), consistent with persistent colocalization of Piwi with Vasa at nuage in early G1/S cells (Figure 3.4G). Likely reflecting the delayed/incomplete translocation of Piwi from mitotic nuage to the nucleus, we also observed that the amount of nuclear Piwi was slightly but significantly reduced in GSCs/SGs in *klp10A*^{RNAi} germ cells. Delayed dissociation of Piwi from mitotic nuage at the end of mitosis was confirmed by live observation of GFP-Piwi and mCherry-Vasa expressed in mitotic SGs. Taken together, these results show that 1) Piwi interacts with nuage specifically in mitosis, and 2) *klp10A* is required for releasing Piwi from mitotic nuage to allow its return to the nucleus at the end of mitosis.

3.5.5 Piwi dissociates from nuage at the central spindle microtubules.

How does Piwi dissociate from mitotic nuage at the end of mitosis, and how does klp10A promote this process? A closer examination of live imaging using GFP-Piwi and mCherry-Vasa revealed that a Piwi-positive compartment ('Piwi granule' hereafter) budded off from mitotic nuage and then dispersed near the center of telophase cells at the central spindle (see below for the confirmation of their localization at the central spindle) (Figure 3.5A). Concomitantly, nuclear Piwi level gradually increased, suggesting that Piwi released from mitotic nuage returned to the nucleus. Morphology of nuage marked by Vasa remained unchanged during this period. In contrast to control SGs, we barely observed the budding off of the Piwi granule in klp10ARNAi SGs (Figure 3.5B). As a result, Piwi remained in nuage for a prolonged time period even after the completion of mitosis. These results show that klp10A is required for releasing Piwi from mitotic nuage

such that Piwi can return to the nucleus. It should be noted that klp10ARNAi germ cells did not exhibit noticeable change in cell cycle progression (see below). Therefore, it is unlikely that prolonged Piwi retention in nuage is an indirect consequence of delayed mitotic exit.

The above live imaging observations indicated that budding off of the Piwi granule from mitotic nuage occurs near the center of dividing cells. Because Klp10A localizes to the central spindle and nuage components likely interact with Klp10A at the central spindle (Figure 3.3), we next examined how budding of the Piwi granule occurs at the central spindle. Using GFP-Piwi combined with mCherry-α-Tubulin, we found that the GFP-Piwi signal was dynamically associated with the central spindle MTs and moved back and forth along MTs, gradually decreasing in intensity until it completely dissipated (Figure 3.5C). In contrast, in klp10ARNAi GSCs/SGs, the majority of Piwi granules did not localize to the central spindle, and even when they are associated with the central spindle, Piwi did not exhibit gradual dispersion as observed in the control (Figure 3.5D).

These results indicate that budding off and release of Piwi from mitotic nuage occurs along the central spindle MTs. We hypothesized that the interaction of nuage with the central spindle MTs facilitates the dissociation of Piwi from nuage. To more directly test this idea that MTs mediate Piwi dissociation from mitotic nuage, we sought to allow cells to exit mitosis in the absence of MTs. To achieve this, we combined ex vivo colcemid treatment with a mad2 mutation. Colcemid treatment effectively depolymerized MTs (Venkei and Yamashita, 2015), which would normally cause metaphase arrest due to the spindle assembly

checkpoint (Li et al., 2010; Venkei and Yamashita, 2015). However, when this is combined with a mad2 mutation, which inactivates the spindle assembly checkpoint, cells exit mitosis even in the absence of MTs (Li et al. 2010). Under this condition, we found that Piwi remained in cytoplasmic nuage after mitotic exit, demonstrating that Piwi's release from mitotic nuage and return to the nucleus is mediated by MTs (Figure 3.5E–G).

3.5.6 klp10A is required for depolymerization of central spindle MTs.

The above results led us to hypothesize that the central spindle serves as a platform for Piwi dissociation from mitotic nuage to allow Piwi to return to the nucleus. How does Klp10A, which is a kinesin motor that bends and depolymerizes MTs (Rogers et al., 2004), regulate this process? Because central spindle MTs seemed to facilitate dissociation of Piwi from mitotic nuage (Figure 3.5), and because Klp10A localizes to the central spindle, we investigated whether Klp10A may regulate the integrity of the central spindle.

To test this idea, we examined the morphology of the central spindle in control vs. klp10ARNAi testes. First, we found that the frequency of GSCs or SGs that contain central spindle MTs was significantly higher in klp10ARNAi compared to control testes (Figure 3.6A–C), suggesting that Klp10A promotes depolymerization of central spindle MTs. In control germ cells, disassembly of the central spindle and completion of S phase, as assessed by pulse labeling with 5-ethynyl-2'-deoxyuridine (EdU), are nearly in synchrony: most of the S phase cells (EdU+) contain central spindle (Figure 3.6D, arrowhead), whereas all post-S phase cells (EdU-) have resolved central spindles (Figure 3.6F). In stark contrast, about half

of post-S phase cells still contained the central spindle in klp10ARNAi testes (Figure 3.6E, arrow, Figure 3.6F), indicating that central spindle disassembly is delayed in the absence of *klp10A*. Importantly, increased frequency of germ cells with the central spindle is not due to skewed composition of cell cycle stages in klp10ARNAi germ cells, because $klp10A^{RNAi}$ germ cells had comparable frequency of being in S phase (22.6±8.6% EdU+ in control, n=89 GSC vs. 24.6±15.6% EdU+ in klp10ARNAi, n=103 GSC, n.s.) or in M phase (0.22±0.03 mitotic GSCs/testis in control, n=160 vs. 0.194±0.01 mitotic GSCs/testis in klp10ARNAi, n=222, n.s.).

Taken together, these results show that *klp10A* is required for depolymerization of central spindle MTs. Based on the observation that Piwi dissociates from mitotic nuage along the central spindle as it disassembles, we speculate that Piwi's dissociation from the central spindle is driven by MT depolymerization, which is facilitated by Klp10A, a MT-depolymerizing kinesin. Understanding how depolymerizing MTs can facilitate dissociation of Piwi from mitotic nuage awaits future investigation.

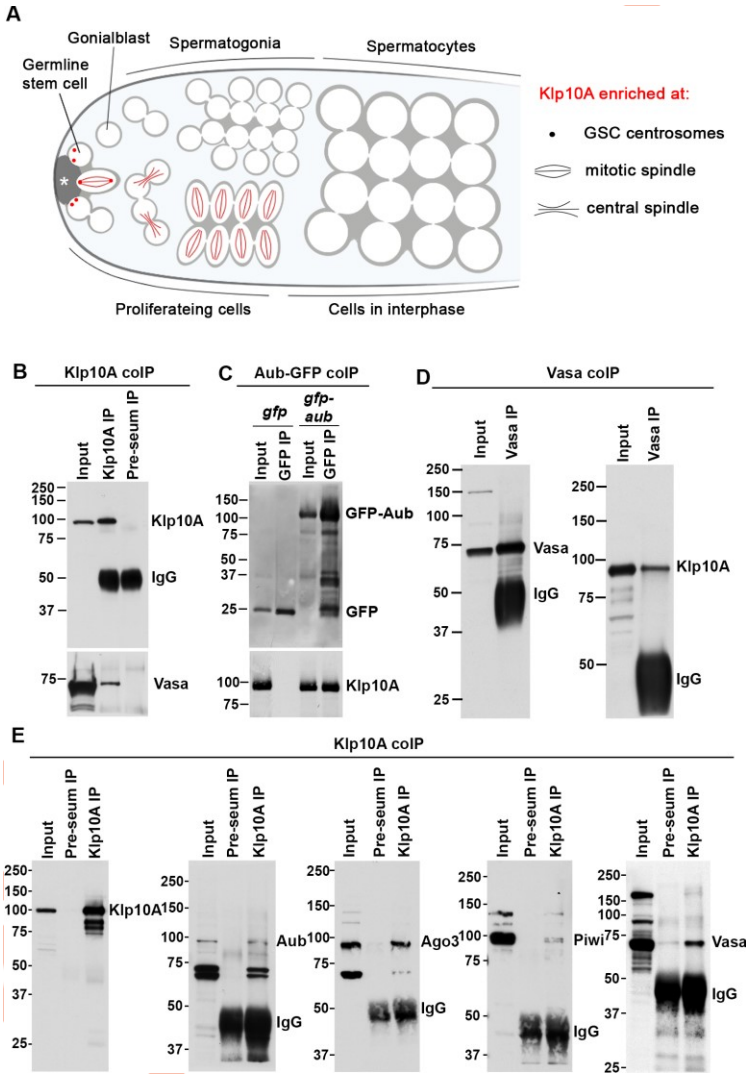

Figure 3.1. Klp10A interacts with the piRNA pathway in GSCs and SGs.

(A) Drosophila spermatogenesis. The stem cell niche is formed by non-dividing somatic cells (hub, marked by asterisk). GSCs are physically attached to the hub, and divide asymmetrically. Gonialblasts (GBs), the differentiating daughters of GSCs, undergo four rounds of mitotic divisions with incomplete cytokinesis. Resultant 16-cell spermatogonia (SGs) then enter meiotic prophase as spermatocytes (SCs). Klp10A is specifically enriched at the centrosomes of GSCs (indicated by red dots), and at the central spindles of GSCs/GBs/SGs (red lines) (Chen et al., 2016). (B) Vasa co-immunoprecipitates with Klp10A in the extract from Upd-expressing testes (nos-gal4>UAS-upd). L: lysate input, IP: immunoprecipitant. (C) Klp10A co-immunoprecipitates with GFP-Aub in the extract from Upd-expressing testes (nos-gal4>UAS-upd UAS-gfp, vs. nos-gal4>UAS-upd UAS-gfp-aub). (D) Klp10A co-immunoprecipites with Vasa in the extract from the Dpp-expressing testes (nos-gal4>UAS-dpp). (E) piRNA pathway

proteins Aub, Ago3, Piwi and Vasa co-immunoprecipitate with Klp10A in the extract from Dpp-expressing testes (nos-gal4>UAS-dpp).

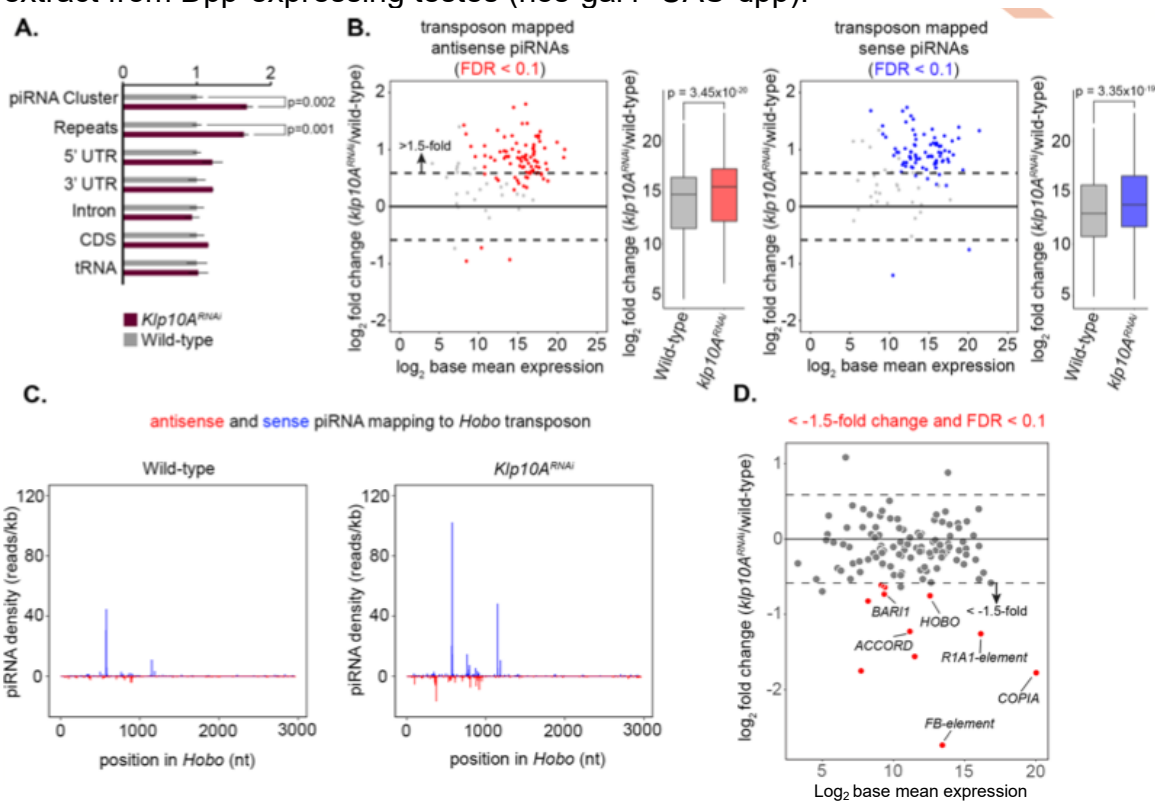

Figure 3.2. piRNA production is upregulated in *klp10A*^{RNAi} testis.

(A) piRNA reads normalized to total library reads at each genomic feature. piRNAs are significantly upregulated in repeat and piRNA cluster regions. Significance was calculated using a Student's t-test. (B) Scatter and bar plots comparing transposon mapped sense and antisense piRNA abundance in *klp10A*^{RNAi} germ cells vs. wild-type. piRNA abundance was calculated by normalizing piRNA read counts mapping to each transposon to total miRNA hairpin reads. A change of greater than 1.5-fold and FDR of <0.1 were used as cutoffs for differential analysis. Significance for piRNA fold change expression was calculated using Wilcoxon signed-rank test. (C) Density of sequenced piRNAs (blue: sense; red: antisense) across *HOBO*. Each piRNA read was normalized to miRNA hairpin abundance. (D) Scatter plot comparing transposon abundance in *klp10A*^{RNAi} versus wild-type (red: FDR<0.1).

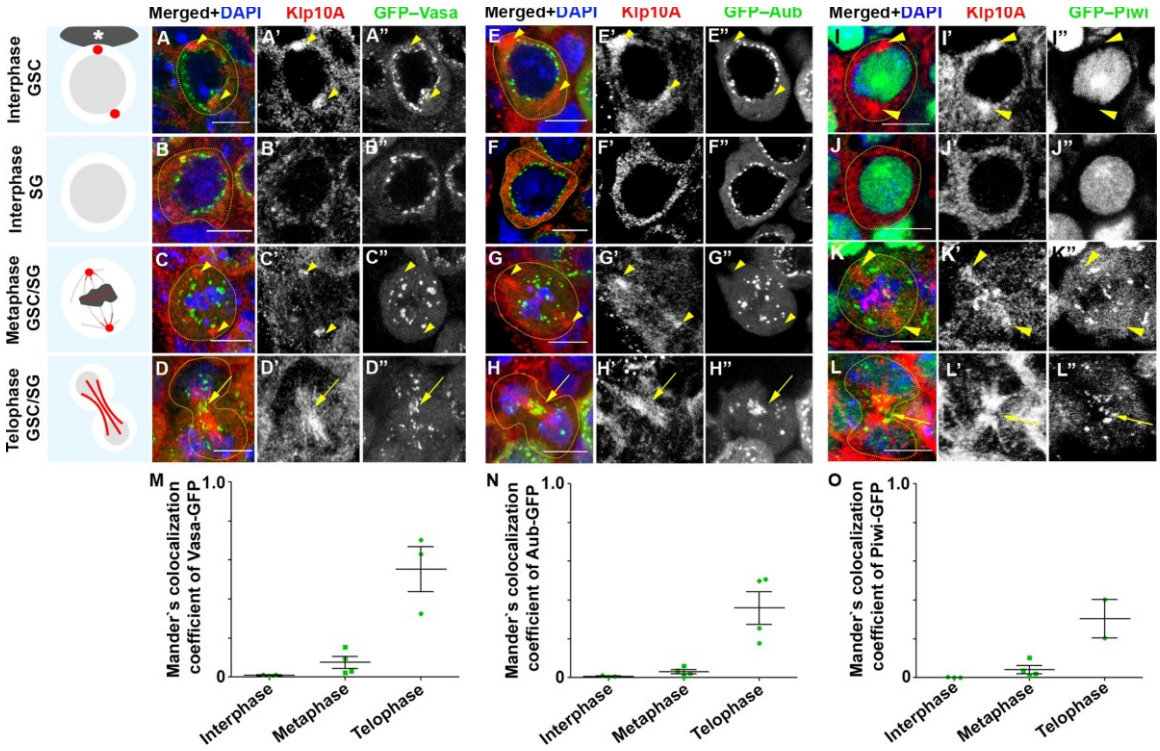

Figure 3.3. Klp10A colocalizes with Vasa, Aub and Piwi at the central spindle in GSCs/SGs during mitotic exit.

A-D'') Localization of Klp10A (red) and GFP-Vasa (green) in GSCs/SGs throughout the cell cycle. DAPI (blue). E-H'') Localization of Klp10A (red) and GFP-Aub (green) in GSCs/SGs throughout the cell cycle, I-L'') Localization of Klp10A (red) and Piwi-GFP (green) in GSCs/SGs throughout the cell cycle. Centrosomal localization of Klp10A is indicated by arrowheads, central spindle localization with arrows. Bars: 5 µm. M-O) Quantification colocalization of GFP-Vasa (M), GFP-Aub (L) and GFP-Piwi (O) with Klp10A in GSCs/SGs during cell cycle. Mender's colocalization coefficient is calculated based on how much of indicated proteins colocalize with Klp10A. Data points represent single cells, each from different testes. Error bars represent SD.

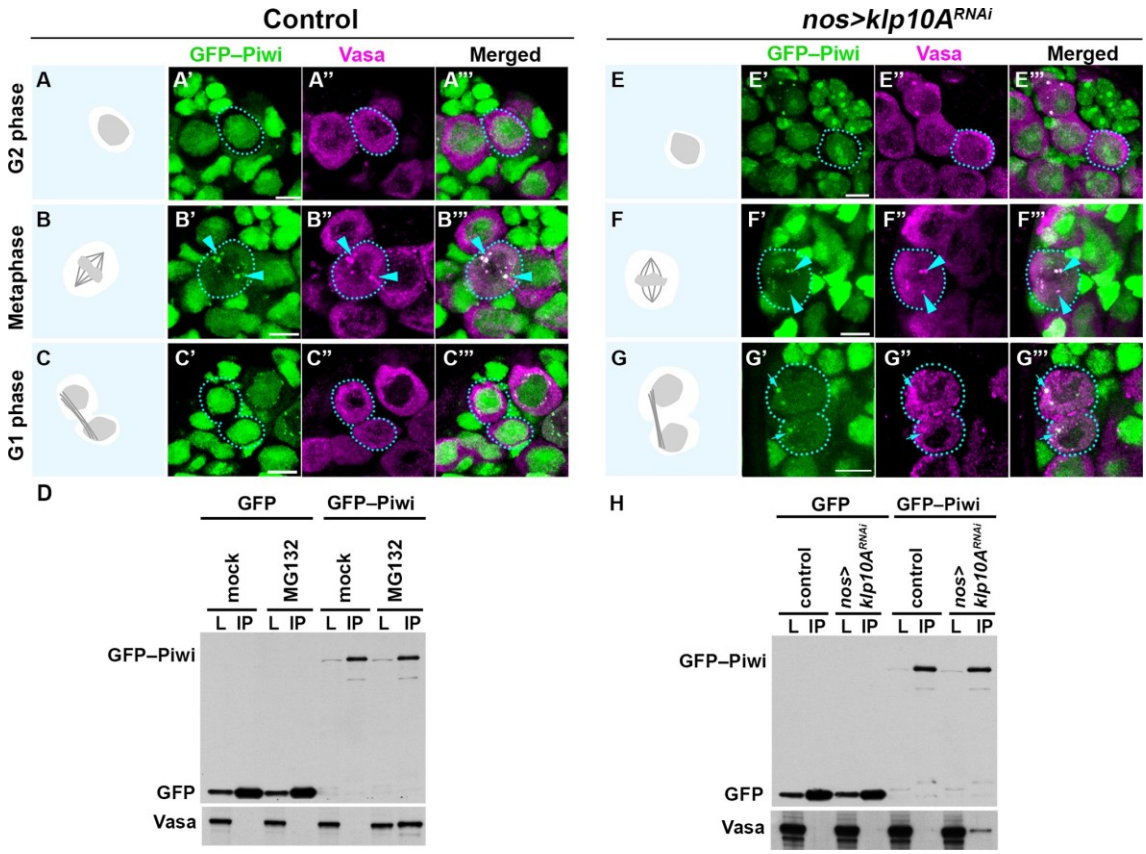

Figure 3.4. Piwi interacts with cytoplasmic piRNA pathway components at the nuage in mitosis, and returns to the nucleus at the end of mitosis in a Klp10A-dependent manner.

(A-C) GFP-Piwi (green) and Vasa (magenta) localization throughout the GSC cell cycle in wild type testes. The localization patterns were the same in GSCs and SGs. GSCs are encircled by dotted lines. Cartoons show DAPI and MT morphology (shown in S8 Fig) to assess cell cycle stages. GFP-Piwi and Vasa colocalization at the mitotic nuage is indicated by arrowheads. Bars: 5 μm. D) co-immunoprecipitation of Vasa with GFP-Piwi from *nos>upd* testes after 4h 30min *ex vivo* MG132 treatment. *nos>upd, gfp* testes were used as control. L: lysate input, IP: immunoprecipitant. (E-G) GFP-Piwi (green) and Vasa (red) localization throughout the GSC cell cycle in *klp10A*^{RNAi} testes. Cytoplasmic GFP-Piwi/Vasa granule in post-mitotic interphase cells are highlighted by arrows in panel G. H) co-immunoprecipitation of Vasa with GFP-Piwi from *nos>upd* testes after *klp10A*^{RNAi}. *nos>upd, gfp* and *nos>upd, gfp, klp10A*^{RNAi} testes were used as control.

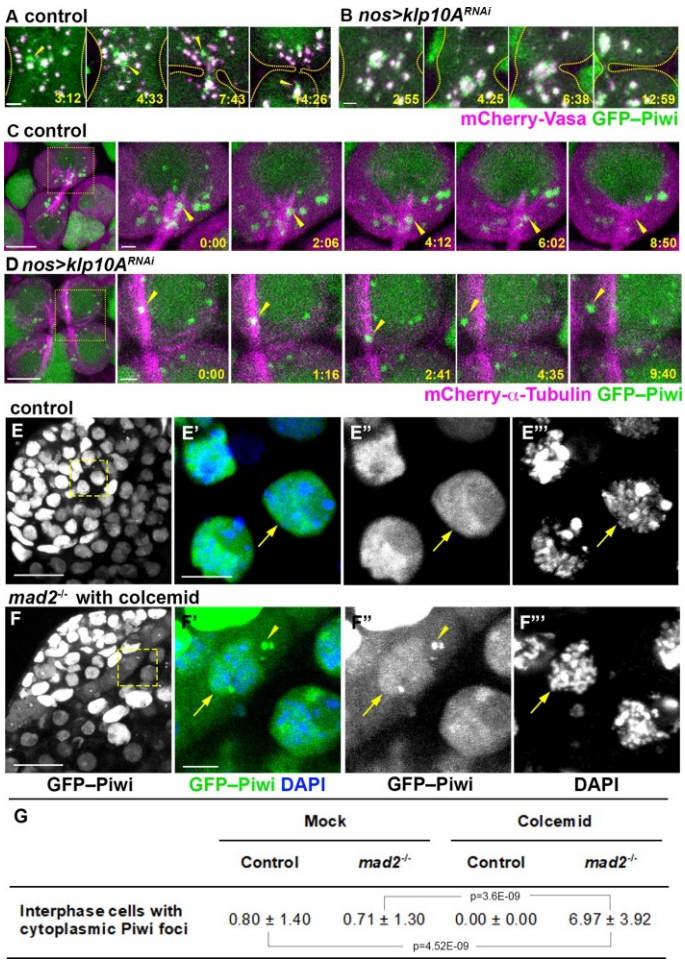

Figure 3.5. Piwi dissociates from mitotic nuage at the central spindle in a *klp10A*-dependent manner.

A-B) Live imaging of mCherry-Vasa (magenta) and GFP-Piwi (green) in the central region of telophase SGs in wild type (A) and in *klp10A*^{RNAi} (B) testes. The elapsed time after anaphase-B onset is indicted in min:sec. Arrowheads point to Piwi-GFP signals budding from the mitotic nuage. Dotted lines outline cells. Bars: 1 µm. C-D) Live imaging of mCherry-α-Tub (magenta) and GFP-Piwi (green) in telophase SGs in wild type (C) and *klp10A*^{RNAi} (D) testes. Squares in the first panels indicate the zoomed regions in the later panels. Arrowheads in (C) point to a mitotic nuage particle, initially positive for Piwi and gradually losing signal. Open arrowheads (D) point to a Piwi-positive nuage particle, sliding along and later releasing the MT bundle, without decreasing GFP-Piwi signal. Bars: 5 µm, and 1 µm for zoomed regions. E-F) GFP-Piwi (green) localization in interphase control SGs (E) or *mad2* null mutant SGs after 3h colcemid treatment (F). Arrowheads point to Piwi-GFP-positive nuage. Arrows point to interphase nuclei. Bars: 20 µm in left panels, 5 µm in insert panels. G) Frequency (cells/per testis) of interphase germ cells with GFP-

Piwi positive nuage 3h after *ex vivo* colcemid treatment. *n*=13-18 testes were used for each experiment.

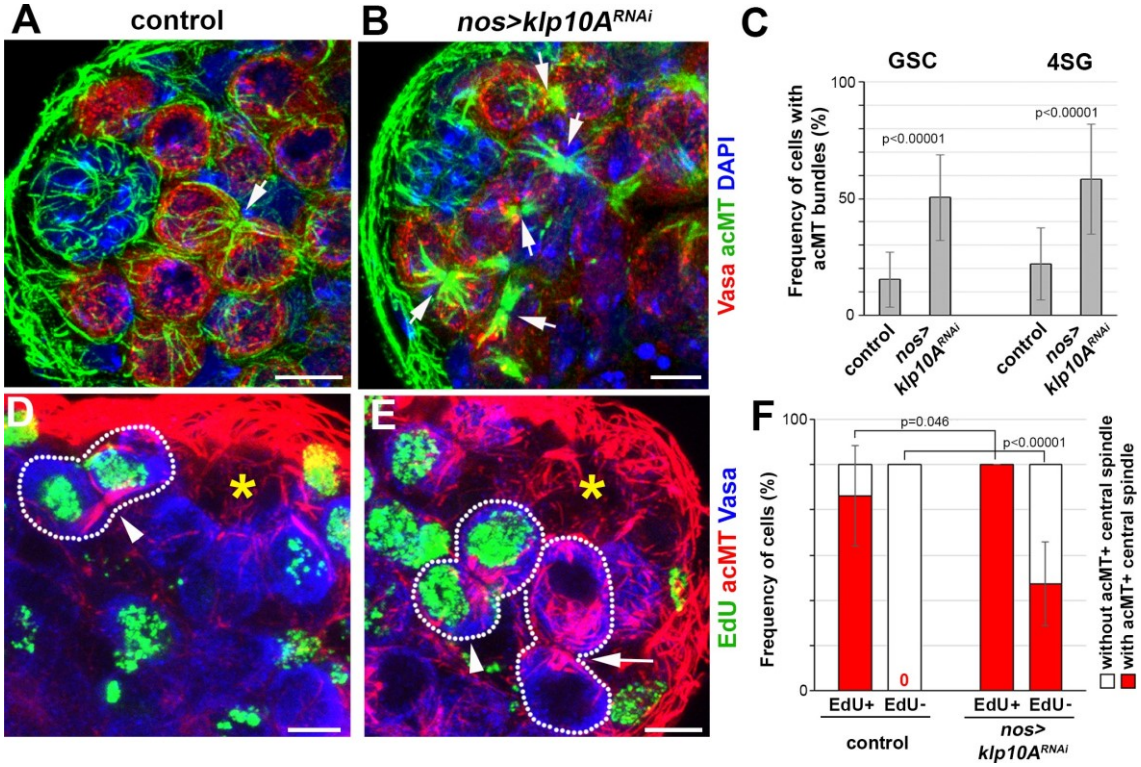

Figure 3.6. Klp10A promotes depolymerization of central spindle MTs.

A, B) Apical tip region of wild type (A) and *nos>klp10A*^{RNAi} (B) testes. Vasa (red), acetylated MTs (acMTs) (green) and DAPI (blue). Arrows indicate acMT bundles of central spindles between interconnected cells. Bars: 5 µm C) Frequency of acMT bundles in GSCs (n=181 for control, n=201 for *klp10A*^{RNAi}) and 4-cell SGs (n=128 and n=95). Error bars indicate SD. p values of t-tests are provided. D-E) Apical tip region of a wild type (D) and a *klp10A*^{RNAi} testis (E) with EdU (green) incorporation after 45min incubation period. Testes were stained for acMTs (red) and Vasa (blue). Arrowheads indicate EdU-positive GSC-GB pairs with acMT bundle. Arrow indicates EdU-negative GSC-GB pairs with acMT bundle in *klp10A*^{RNAi}. Asterisks indicate the hub, dotted lines indicate GSC-GB pairs. Bars: 5 µm. F) Frequency of acMT bundle in EdU-positive (n=32 for control, n=24 for *klp10A*^{RNAi}) and EdU-negative GSCs (n=69 for control, n=79 for *klp10A*^{RNAi}). Error bars indicate SD. p values of t-tests are provided.

3.6 Discussion

In this study, we reveal an unexpected, dynamic localization of piRNA pathway machinery during the cell cycle of mitotically proliferating germ cells (GSCs and SGs) in the *Drosophila* testis. Our study is the first to describe a dynamic compositional change in nuage during the cell cycle. Piwi is nuclear in interphase, but associates with nuage specifically during mitosis and then returns to the nucleus at the end of the mitosis. Our study provides several novel insights into the cell biological aspects of piRNA biogenesis.

Our data suggest that this 'nuage cycle' in mitotic germ cells of the testis is partly facilitated by disassembly of central spindle MTs: Klp10A regulates depolymerization of central spindle MTs during telophase, and Piwi dissociates from nuage at depolymerizing central spindle MTs. How can depolymerizing MTs facilitate dissociation of Piwi from the nuage compartment? Although the underlying molecular mechanisms remain elusive, the physical interaction of piRNA pathway components and MTs have also been documented (Lau et al., 2009; Rodriguez et al., 2005; Sato et al., 2011). In addition, MTs have been shown to regulate the assembly, maturation, and disassembly of stress granules, another phase-separated compartment similar to nuage (Chernov et al., 2009; Ivanov et al., 2003; Loschi et al., 2009; Shao et al., 2017). Given these earlier studies, our observation might reflect a generalizable regulatory feature of the piRNA pathway. Recently, it was shown that liquid droplets of Tau protein concentrate tubulin dimers to facilitate MT formation, and that MTs within the Tau droplets exhibit liquid-like properties (Hernández-Vega et al., 2017). Polymerized MTs within Tau

droplets then deforms Tau droplets (Hernández-Vega et al., 2017). Likewise, tubulin-binding proteins within nuage (Lau et al., 2009; Rodriguez et al., 2005; Sato et al., 2011) may increase tubulin concentration, leading to MT polymerization. It might in turn deform nuage and change biophysical characteristics of phase separation, leading to release of Piwi from nuage. Although our results show the role of Klp10A in dissociating Piwi from nuage at the end of mitotic spindle, underlying mechanism of how Klp10A interacts with piRNA pathway components in facilitating Piwi's dissociation on the depolymerizing MTs remains unknown. Klp10A protein does not have any noticeable domains other than central motor domain, and the structure does not hint as to how Klp10A might interact with piRNA pathway components. It is possible that Klp10A's interaction with nuage components are mediated by unidentified linker protein(s). Also, it remains unknown how Piwi associates with nuage at the mitotic entry, which is likely independent of Klp10A. Nuclear envelope breakdown may be sufficient for Piwi to associate with nuage. Alternatively, there may be additional post-translational regulations that enhance Piwi's affinity to nuage.

We demonstrate that Klp10A is involved in piRNA biogenesis. Loss of klp10A leads to a global upregulation of piRNA expression but only modest depletion of a subset of transposons. The failure of *klp10A* knockdown to show depletion of a majority of TEs, despite elevated levels of cognate piRNAs, could be because *klp10A* influences piRNA biogenesis only during limited stages of spermatogenesis, i.e. GSCs and SGs. For example, TE transcripts are not regulated by the piRNA pathway in germ cells of certain stages of

spermatogenesis (Vourekas et al., 2012), and defects caused by klp10A depletion may be masked by unchanged TE expression in other stages of spermatogenesis. Our cytological data show that klp10A is required for the release of Piwi from the mitotic nuage, such that Piwi can re-enter the nucleus upon completion of mitosis. As a result of delayed return of Piwi to the nucleus, overall nuclear amount of Piwi was reduced in *klp10A*^{RNAi}. How do these defects result in increased piRNA production? It is possible that the prolonged localization of Piwi to the nuage in *klp10A*^{RNAi} may enhance the production of pre-piRNAs (or its loading to Piwi), leading to subsequent dysregulation of ping-pong activity to cause upregulation of piRNAs. Alternatively, because Piwi represses the expression of TEs and certain pre-piRNAs (Chang et al., 2019; Sato and Siomi, 2018), reduced nuclear Piwi might lead to derepression of TEs and pre-piRNAs, which, in turn, may result in enhanced piRNA production through intact ping-pong activity in *klp10A*^{RNAi} SGs.

Expression of piRNAs in male germ cells appears to be distinct from those found in nurse cells during female germline development (Malone and Hannon, 2009). Nurse cells are highly polyploid post-mitotic cells and thus do not undergo nuclear envelope breakdown as do SGs (Spradling, 1993). In nurse cells of the *Drosophila* ovary, Piwi needs to be loaded with piRNAs to translocate into the nucleus (Handler et al., 2011; Olivieri et al., 2010; Wang et al., 2015). However, Piwi localization to nuage is not visible in wild type cells, likely because Piwi only transiently associates with nuage prior to nuclear translocation. Only when the piRNA pathway is completely compromised, as in the aub ago3 double mutant, can the failure of Piwi to enter the nucleus be detected, leading to its accumulation

in nuage (Wang et al., 2015). In contrast to this mechanism of piRNA loading and nuclear translocation of Piwi in post-mitotic nurse cells, our finding suggests that mitotically dividing germ cells (SGs) in the testis may utilize mitosis (nuclear envelope breakdown) as a means to load Piwi with piRNAs. Nuclear envelope breakdown may represent a robust and efficient way of loading Piwi with piRNAs in SGs that divide every ~12 hours.

Taken together, we propose that Klp10A-dependent MT depolymerization at the central spindle facilitates Piwi dissociation from nuage to promote its translocation back into the nucleus. Piwi's interaction with nuage during mitosis might impact piRNA biogenesis, as indicated by piRNA profiling. In summary, the present study reveals a novel cell biological mechanism by which nuclear and cytoplasmic arms of piRNA biogenesis communicate during the cell cycle.

3.7 Materials and methods

Fly husbandry and transgenic flies

Flies were raised in standard Bloomington medium at 25°C. The following stocks were used. UAS-gfp-klp10A (Inaba et al., 2015a), gfp-vas (Sano et al., 2002), gfp-piwi (a gift from Katalin Fejes Tóth) (Le Thomas et al., 2013), mCherry-vas (a gift from Elizabeth Gavis)(Lerit and Gavis, 2011), and the following stocks, obtained from the Bloomington Stock Center: nos-gal4 (Van Doren et al., 1998), UAS-upd (Zeidler et al., 1999), UAS-gfp (Spana E., 1999, personal communication to FlyBase, ID: FBrf0111645), UAS-gfp-aub (Harris and Macdonald, 2001), UAS-mCherry-α-tubulin (Rusan and Peifer, 2007), UAS-klp10ATRiP.HMS00920

(Flybase: FBrf0214641, FBrf0212437), mad2EY21687 (Li et al., 2010), Df(3L)BSC437 (FlyBase: FBrf0204472)

Immunoprecipitation and western blotting

For immunoprecipitation, 150 pairs of testes were lysed in 0.5 ml HEPES based buffer [50mM HEPES pH 7.5, 250mM NaCl, 0.1% Nonidet P40, 0.2% Triton X-100, supplied with cOmplete protease inhibitor mixture (Roche)]. In Klp10A co-immunoprecipitation (co-IP) experiments, the cleared lysates were incubated with rabbit anti-Klp10A serum (1:200, (Chen et al., 2016)) for 4 hours, extended by 2 extra hours after supplement with Protein A Dynabeads (Life Technologies). The beads were washed four times with lysis buffer and proteins were eluted with SDS-PAGE protein sample buffer. For co-IP experiments with GFP tagged proteins, the cleared lysates were incubated with GFP-Trap magnetic beads (Chromotek) according to manufacturer's instructions. To detect precipitated proteins on western-blots, the following primary antibodies were combined with horseradish peroxidase conjugated secondary antibodies (Abcam): guinea-pig anti-Klp10A (1:10,000, (Chen et al., 2016)), rabbit anti-GFP (1:3000, ab290, Abcam), rabbit anti-Vasa d-26 (1:3000, Santa Cruz Biotechnology), rabbit anti-Piwi (1:3000 (Klattenhoff et al., 2009)), rabbit anti-Aub (1:3000 (Klattenhoff et al., 2009)), and mouse anti-Ago3 (1:3000 (Senti et al., 2015)).

Mass-spectrometry

Klp10A was purified either with anti-Klp10A or anti-GFP nanobodies from GSC-enriched testes (250 pairs of *nos-gal4>UAS-upd* testes or 1500 pairs of nos-gal4>UAS-upd, UAS-gfp-klp10A testes, respectively) were dissected, and

homogenized in lysis buffer (10 mM Tris-HCl pH 7.5; 150 mM NaCl; 0.5 mM EDTA, 0.1% NP40 and protease inhibitor cocktail (cOmplete, Roche)) for 30 min at 4°C. After centrifugation at 13,000 rpm for 10-15 min at 4°C the supernatant was saved as whole cell extract. The protein concentration was measured by absorbance at 562 nm using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). The whole cell extract of 1 µg total protein mass was incubated with 40 µl packed guinea-pig anti-Klp10A (Chen et al., 2016) conjugated Protein A-Dynabeads (Thermo Fisher Scientific) for 2h at 4°C, or with 25 µl packed GFP-Trap agarose beads (Chromotek) for 3 hours at 4°C. Beads were washed three times with wash buffer (10 mM Tris-HCl pH 7.5; 150 mM NaCl; 0.5 mM EDTA), and the proteins were eluted in LDS sample loading buffer (1.5x) at 100˚C for 5 min and separated on a 10% Bis-Tris Novex mini-gel (Invitrogen) using the MES buffer system. In the anti-Klp10A pull down experiments the gel was stained with Coomassie dye and excised into ten equally sized segments. These segments were analyzed by LC/MS/MS (MS Bioworks, Ann Arbor, MI). The gel digests were analyzed by nano LC/MS/MS with a Waters NanoAcquity HPLC system interfaced to a Thermo Fisher Q Exactive. Peptides were loaded on a trapping column and eluted over a 75 µm analytical column at 350 nL/min; both columns were packed with Luna C18 resin (Phenomenex). The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed in the Orbitrap at 70,000 FWHM resolution and 17,500 FWHM resolution, respectively. In the anti-GFP pull down experiments after electrophoresis the gel was stained with coomassie and excised into two equally sized segments. The gel digests were analyzed by LC-MS/MS with LTQ

Orbitrap XL mass spectrometer (Thermo Fisher Scientific) at Fred Hutchinson Cancer Research Center, Proteomics Resource.

Immunofluorescence staining and microscopy

Fixation and immunofluorescence staining of testes was performed as described previously (Chen et al., 2016). The following primary antibodies were used. Mouse anti-Fasciclin III (FasIII; 1:100; 7G10, developed by Goodman C, obtained from Developmental Studies Hybridoma Bank; (Patel et al., 1987)), rabbit anti-Vasa (1:200; d-26, Santa Cruz Biotechnology), mouse anti-□-tubulin (4.3; 1:50; developed by C. Walsh (Walsh, 1984) and obtained from the Developmental Studies Hybridoma Bank), mouse anti-acetylated-Tubulin (1:100, 6-11B-1, Sigma), rabbit anti-phosphorylated (Thr3) histone H3 (PH3) (1:200; clone JY325, Upstate), mouse anti-LaminB (1:200; C20; Santa Cruz Biotechnology), guinea pig anti-Traffic jam (Tj; 1:400; a gift from Dorothea Godt; (Li et al., 2003)), rabbit and guinea-pig anti-Klp10A (1:3000/1:1000 respectively; (Chen et al., 2016)). Alexa Fluor-conjugated secondary antibodies were used (1:200; Thermo Fisher Scientific), EdU was detected by Click-iT Plus EdU Imaging Kit with Alexa Fluor 647 (Thermofisher). Testes were mounted into VECTASHIELD media with 4',6-diamidino-2-phenylindole (DAPI; Vector Labs). Images were captured using a Leica TCS SP8 confocal microscope with a 63×oil-immersion objective (NA=1.4) and processed by ImageJ software (Schindelin et al., 2012). Colocalization of nuage components with Klp10A or acMTs were quantified by calculating Mander`s coefficient (Adler and Parmryd, 2010; Dunn et al., 2011) in ImageJ.

***Ex vivo* treatment of *Drosophila* testis**

To enrich cells in metaphase testes were dissected and transferred to Schneider's insect medium (Sigma) containing MG132 (20 μM final concentration, Sigma-Aldrich) (Moutinho-Pereira et al., 2010) or colcemid (100μM final concentration, Calbiochem) (Venkei and Yamashita, 2015). After 4.5h incubation with MG132 or colcemid at 25°C, GSCs/SGs arrested in metaphase with or without intact bipolar mitotic spindles, respectively. For EdU incorporation the dissected testes were incubated for 45min (10 μM, Thermofisher) in Schneider's insect medium (Gibco) at 25°C.

RNA extraction, library preparations, and sequencing

Total RNAs were extracted by Trizol. 300 pairs of testes were used for each sample, and biological triplicates for each genotype were used. For small RNA sequencing, 2.5 μg of total RNAs were treated with DNase I (Amplification Grade, ThermoFisher) and recovered by RNA Clean & Concentrator-5 columns (Zymo). Libraries were generated from DNaseI treated RNAs using reagents from the TruSeq Small RNA Library Prep Kit (Illumina), while following a previously published protocol from (Wickersheim and Blumenstiel, 2013) that uses a terminator oligo to deplete *Drosophila* 2S rRNA from the final sequencing libraries. The libraries were sequenced in one lane on a HiSeq 2500 machine (Illumina) at the UCLA Broad Stem Cell Research Center BioSequencing Core. The terminator oligo complimentary to *Drosophila* 2S rRNA was TAC AAC CCT CAA CCA TAT GTA GTC CAA GCA /3SpC3/ (Integrated DNA Technologies).

For mRNA sequencing, 2.5 μg of total RNAs were treated with DNase I (Amplification Grade, ThermoFisher) and recovered by RNA Clean &

Concentrator-25 columns (Zymo). DNaseI treated RNAs were then treated with Ribo-Zero Gold rRNA Removal Kit (Illumina). Libraries were generated from rRNA depleted RNAs using TruSeq Stranded Total RNA Sample Prep Kit (Illumina). The libraries were sequenced by a HiSeq 4000 machine (Illumina) at the UCLA Broad Stem Cell Research Center BioSequencing Core.

RNA-seq data analysis

Three biological replicates of each condition were used to test for statistical significance and comparative analysis. Raw qseq files were converted to fastq format using custom Python and awk scripts. Small RNAs were clipped from 3' adapter sequences using Trimmomatic 0.27 (Bolger et al., 2014). Reads mapping with at most 1 mismatch to rRNA and miRNA hairpin sequences were parsed out by Bowtie2 v2.2.3 (Langmead and Salzberg, 2012). The remaining reads were aligned to the *D. melanogaster* transcriptome (Dm3) with at most 1 mismatch allowed using Bowtie2 v2.2.3. Abundance estimation was done at annotated genes, piRNA clusters, and TEs using the eXpress pipeline within the piPipes suite (Han et al., 2015; Roberts and Pachter, 2012). We used Flybase for annotation of protein coding genes and TEs. piRNA cluster annotations were done according to (Brennecke et al., 2007). For piRNA analysis, we selected reads that were 23-29 nucleotides in length to ensure other endo-siRNA species did not influence downstream analysis. miRNA hairpin reads were used to normalize between libraries. Differential expression analysis was done in DESeq2 3.7 using the Wald test and adjusted p-values were corrected using the Benjamin-Hochberg

procedure with an FDR threshold of 0.01(Love et al., 2014). Plots were generated using R. Sequencing data is available at GEO: accession number: GSE122596.

Ping-pong activity analysis:

piRNA reads were mapped directly to the *D. melanogaster* transcriptome (Dm3) using Bowtie2 v2.2.3 (Langmead and Salzberg, 2012). BAM files from Bowtie2 were converted to BED files using several pipelines, including BEDTools 2.7 from the piPipes suite, which assigns nucleotide positions of each read across TE transcript bodies (Han et al., 2015; Quinlan and Hall, 2010). In addition to nucleotide position, reads were categorized as sense and antisense. Sense was defined as piRNAs that were derived from cleavage of the TE mRNAs and antisense was defined as piRNAs that were antisense to annotated TE transcripts. Custom Python scripts were used to calculate the number of 5' to 5' complementarity between sense and antisense reads. Ping-pong ratios were calculated for each transposon feature across all libraries. To calculate the ping-pong ratio of each transposon, we used custom Python scripts to calculate the number of piRNAs in which sense piRNAs with an A at the 10-nt position or antisense piRNAs with a U at the 1-nt position showed 10 nt of complementarity from the 5' end. We then divided the number of such pairs by the total number of piRNA reads. The resulting ratio allowed for quantification of ping-pong activity, without needing to normalize for library size. The calculation of ping-pong ratio was done using custom Python scripts. Plots were generated using R.

mRNA-seq data analysis

Raw qseq files were converted to fastq format using custom Python and awk scripts. Due to the repetitive nature of transposon sequences, reads were aligned to the *D. melanogaster* genome (*Dm3*) by STAR-2.6.0a while allowing for up to 100 multi-alignments per read (Dobin et al., 2013). Feature and abundance estimation were determined using TEtranscripts 2.0.3 (Jin et al., 2015). TEtranscripts initially distributes multi-mapped reads evenly among potential matches and optimizes the distribution of multi-mapped reads using the Expectation Maximization approach. Variance was measured based on normalized gene expression counts and we performed principal components analysis and found samples derived from wild-type clustered separately from klp10ARNAi on Principle Component 1 (S6B Fig). Differential analysis was done using DESeq2 3.7 using the Wald test and p-values were corrected using the Benjamin-Hochberg procedure with an FDR threshold of 0.01 (Love et al., 2014). All Plots were generated using R. Pipeline of Analysis is described in (S1A and B Fig).

Live imaging

Testes from newly enclosed flies were dissected in Schneider's Drosophila medium (Gibco) and prepared for live imaging as described previously (Cheng and Hunt, 2009). The testis tips were placed into a drop of medium in a glass-bottom chamber and were covered by regenerated cellulose membrane (Spectrum Lab). The chamber was mounted on a three-axis computer-controlled piezoelectric stage. An inverted Leica TCS SP8 confocal microscope with a 63× oil immersion

objective (NA = 1.4) was used for imaging. Live imaging was performed at ambient room temperature. Images were processed using ImageJ software.

3.7 Acknowledgements

We thank Elizabeth Gavis, Katalin Fejes Tóth, Dorothea Godt, Mikiko Siomi, Phil Zamore, Bill Theurkauf, the Bloomington Stock Center and the Developmental Studies Hybridoma Bank for reagents, the Yamashita lab members, Rebecca Tay, Margaret Starostik, Nelson Lau, Sue Hammoud and Lei Lei for discussions and comments on the manuscript. This research was supported by Howard Hughes Medical Institute (to Y.Y. and S.E.J.) and NIH R01GM118875 (to J.K.K.).

3.8 References

Adler, J., and I. Parmryd. 2010. Quantifying colocalization by correlation: The Pearson correlation coefficient is superior to the Mander's overlap coefficient. Cytometry Part A. 77A:733-742.

Andress, A., Y. Bei, B.R. Fonslow, R. Giri, Y. Wu, J.R. Yates, 3rd, and R.W. Carthew. 2016. Spindle-E cycling between nuage and cytoplasm is controlled by Qin and PIWI proteins. J Cell Biol. 213:201-211.

Aravin, A.A., M. Lagos-Quintana, A. Yalcin, M. Zavolan, D. Marks, B. Snyder, T. Gaasterland, J. Meyer, and T. Tuschl. 2003. The small RNA profile during *Drosophila melanogaster* development. Dev Cell. 5:337-350.

Aravin, A.A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, K.F. Toth, T. Bestor, and G.J. Hannon. 2008. A piRNA pathway primed by individual

transposons is linked to de novo DNA methylation in mice. Mol Cell. 31:785-799.

Batista, P.J., J.G. Ruby, J.M. Claycomb, R. Chiang, N. Fahlgren, K.D. Kasschau, D.A. Chaves, W. Gu, J.J. Vasale, S. Duan, D. Conte, Jr., S. Luo, G.P. Schroth, J.C. Carrington, D.P. Bartel, and C.C. Mello. 2008. PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. Mol Cell. 31:67-78.

Bolger, A.M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114-2120.

Brangwynne, C.P., C.R. Eckmann, D.S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Julicher, and A.A. Hyman. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science. 324:1729-1732.

Brennecke, J., A.A. Aravin, A. Stark, M. Dus, M. Kellis, R. Sachidanandam, and G.J. Hannon. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell. 128:1089-1103.

Chang, T.H., E. Mattei, I. Gainetdinov, C. Colpan, Z. Weng, and P.D. Zamore. 2019. Maelstrom Represses Canonical Polymerase II Transcription within Bi-directional piRNA Clusters in *Drosophila melanogaster*. Molecular Cell. 73:291-303.e296.

Chen, C., M. Inaba, Z.G. Venkei, and Y.M. Yamashita. 2016. Klp10A, a stem cell centrosome-enriched kinesin, balances asymmetries in Drosophila male germline stem cell division. Elife. 5.

Cheng, J., and A.J. Hunt. 2009. Time-lapse live imaging of stem cells in Drosophila testis. Curr Protoc Stem Cell Biol. Chapter 2:Unit 2E 2.

Cheng, J., N. Turkel, N. Hemati, M.T. Fuller, A.J. Hunt, and Y.M. Yamashita. 2008. Centrosome misorientation reduces stem cell division during ageing. Nature. 456:599-604.

Chernov, K.G., A. Barbet, L. Hamon, L.P. Ovchinnikov, P.A. Curmi, and D. Pastré. 2009. Role of Microtubules in Stress Granule Assembly: MICROTUBULE DYNAMICAL INSTABILITY FAVORS THE FORMATION OF MICROMETRIC STRESS GRANULES IN CELLS. Journal of Biological Chemistry. 284:36569-36580.

Conduit, P.T., and J.W. Raff. 2010. Cnn dynamics drive centrosome size asymmetry to ensure daughter centriole retention in drosophila neuroblasts. Curr Biol. 20:2187-2192.

Cox, D.N., A. Chao, and H. Lin. 2000. piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells. Development. 127:503-514.

Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T.R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29:15-21.

Dunn, K.W., M.M. Kamocka, and J.H. McDonald. 2011. A practical guide to evaluating colocalization in biological microscopy. American Journal of Physiology-Cell Physiology. 300:C723-C742.

Girard, A., R. Sachidanandam, G.J. Hannon, and M.A. Carmell. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature. 442:199-202.

Gonzalez, J., H. Qi, N. Liu, and H. Lin. 2015. Piwi Is a Key Regulator of Both Somatic and Germline Stem Cells in the Drosophila Testis. Cell Rep. 12:150-161.

Gunawardane, L.S., K. Saito, K.M. Nishida, K. Miyoshi, Y. Kawamura, T. Nagami, H. Siomi, and M.C. Siomi. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science. 315:1587-1590.

Han, B.W., W. Wang, P.D. Zamore, and Z. Weng. 2015. piPipes: a set of pipelines for piRNA and transposon analysis via small RNA-seq, RNA-seq, degradome- and CAGE-seq, ChIP-seq and genomic DNA sequencing. Bioinformatics. 31:593-595.

Handler, D., D. Olivieri, M. Novatchkova, F.S. Gruber, K. Meixner, K. Mechtler, A. Stark, R. Sachidanandam, and J. Brennecke. 2011. A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. The EMBO journal. 30:3977-3993.

Harris, A.N., and P.M. Macdonald. 2001. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. Development. 128:2823-2832.

Hernández-Vega, A., M. Braun, L. Scharrel, M. Jahnel, S. Wegmann, B.T. Hyman, S. Alberti, S. Diez, and A.A. Hyman. 2017. Local Nucleation of Microtubule

Bundles through Tubulin Concentration into a Condensed Tau Phase. Cell Rep. 20:2304-2312.

Inaba, M., M. Buszczak, and Y.M. Yamashita. 2015a. Nanotubes mediate niche-stem-cell signalling in the Drosophila testis. Nature. 523:329-332.

Inaba, M., Z.G. Venkei, and Y.M. Yamashita. 2015b. The polarity protein Baz forms a platform for the centrosome orientation during asymmetric stem cell division in the Drosophila male germline. Elife. 4.

Ivanov, P.A., E.M. Chudinova, and E.S. Nadezhdina. 2003. Disruption of microtubules inhibits cytoplasmic ribonucleoprotein stress granule formation. Experimental Cell Research. 290:227-233.

Januschke, J., S. Llamazares, J. Reina, and C. Gonzalez. 2011. Drosophila neuroblasts retain the daughter centrosome. Nat Commun. 2:243.

Januschke, J., J. Reina, S. Llamazares, T. Bertran, F. Rossi, J. Roig, and C. Gonzalez. 2013. Centrobin controls mother-daughter centriole asymmetry in Drosophila neuroblasts. Nature cell biology. 15:241-248.

Jin, Y., O.H. Tam, E. Paniagua, and M. Hammell. 2015. TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics. 31:3593-3599.

Kawase, E., M.D. Wong, B.C. Ding, and T. Xie. 2004. Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the Drosophila testis. Development. 131:1365-1375.

Kibanov, M.V., K.S. Egorova, S.S. Ryazansky, O.A. Sokolova, A.A. Kotov, O.M. Olenkina, A.D. Stolyarenko, V.A. Gvozdev, and L.V. Olenina. 2011. A novel

organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing. Mol Biol Cell. 22:3410-3419.

Kiger, A.A., D.L. Jones, C. Schulz, M.B. Rogers, and M.T. Fuller. 2001. Stem cell self-renewal specified by JAK-STAT activation in response to a support cell cue. Science. 294:2542-2545.

Klattenhoff, C., H. Xi, C. Li, S. Lee, J. Xu, J.S. Khurana, F. Zhang, N. Schultz, B.S. Koppetsch, A. Nowosielska, H. Seitz, P.D. Zamore, Z. Weng, and W.E. Theurkauf. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell. 138:1137-1149.

Klenov, M.S., S.A. Lavrov, A.P. Korbut, A.D. Stolyarenko, E.Y. Yakushev, M. Reuter, R.S. Pillai, and V.A. Gvozdev. 2014. Impact of nuclear Piwi elimination on chromatin state in Drosophila *melanogaster* ovaries. Nucleic Acids Res. 42:6208-6218.

Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods. 9:357.

Lau, N.C., T. Ohsumi, M. Borowsky, R.E. Kingston, and M.D. Blower. 2009. Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. The EMBO Journal. 28:2945.

Le Thomas, A., A.K. Rogers, A. Webster, G.K. Marinov, S.E. Liao, E.M. Perkins, J.K. Hur, A.A. Aravin, and K.F. Toth. 2013. Piwi induces piRNA-guided

transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. 27:390-399.

Lerit, D.A., and E.R. Gavis. 2011. Transport of germ plasm on astral microtubules directs germ cell development in Drosophila. Current biology : CB. 21:439-448.

Li, C., V.V. Vagin, S. Lee, J. Xu, S. Ma, H. Xi, H. Seitz, M.D. Horwich, M. Syrzycka, B.M. Honda, E.L. Kittler, M.L. Zapp, C. Klattenhoff, N. Schulz, W.E. Theurkauf, Z. Weng, and P.D. Zamore. 2009. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell. 137:509-521.

Li, D., G. Morley, M. Whitaker, and J.Y. Huang. 2010. Recruitment of Cdc20 to the kinetochore requires BubR1 but not Mad2 in Drosophila *melanogaster*. Mol Cell Biol. 30:3384-3395.

Li, M.A., J.D. Alls, R.M. Avancini, K. Koo, and D. Godt. 2003. The large Maf factor Traffic Jam controls gonad morphogenesis in Drosophila. Nat Cell Biol. 5:994-1000.

Lim, A.K., and T. Kai. 2007. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila *melanogaster*. Proc Natl Acad Sci U S A. 104:6714-6719.

Loschi, M., C.C. Leishman, N. Berardone, and G.L. Boccaccio. 2009. Dynein and kinesin regulate stress-granule and P-body dynamics. Journal of cell science. 122:3973-3982.

Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 15:550.

Malone, C.D., J. Brennecke, M. Dus, A. Stark, W.R. McCombie, R. Sachidanandam, and G.J. Hannon. 2009. Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell. 137:522-535.

Malone, C.D., and G.J. Hannon. 2009. Molecular Evolution of piRNA and Transposon Control Pathways in Drosophila. Cold Spring Harbor Symposia on Quantitative Biology. 74:225-234.

Moutinho-Pereira, S., I. Matos, and H. Maiato. 2010. Drosophila S2 cells as a model system to investigate mitotic spindle dynamics, architecture, and function. Methods Cell Biol. 97:243-257.

Nagao, A., T. Mituyama, H. Huang, D. Chen, M.C. Siomi, and H. Siomi. 2010. Biogenesis pathways of piRNAs loaded onto AGO3 in the Drosophila testis. RNA.

Nishida, K.M., Y.W. Iwasaki, Y. Murota, A. Nagao, T. Mannen, Y. Kato, H. Siomi, and M.C. Siomi. 2015. Respective functions of two distinct Siwi complexes assembled during PIWI-interacting RNA biogenesis in Bombyx germ cells. Cell Rep. 10:193-203.

Nishida, K.M., K. Saito, T. Mori, Y. Kawamura, T. Nagami-Okada, S. Inagaki, H. Siomi, and M.C. Siomi. 2007. Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad. RNA. 13:1911-1922.

Nott, T.J., E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T.D. Craggs, D.P. Bazett-Jones, T. Pawson, J.D. Forman-Kay, and A.J. Baldwin. 2015. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Mol Cell. 57:936-947.

Olivieri, D., M.M. Sykora, R. Sachidanandam, K. Mechtler, and J. Brennecke. 2010. An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. EMBO J. 29:3301-3317.

Patel, N.H., P.M. Snow, and C.S. Goodman. 1987. Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in Drosophila. Cell. 48:975-988.

Pek, J.W., and T. Kai. 2011. A role for vasa in regulating mitotic chromosome condensation in Drosophila. Current biology : CB. 21:39-44.

Quinlan, A.R., and I.M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841-842.

Roberts, A., and L. Pachter. 2012. Streaming fragment assignment for real-time analysis of sequencing experiments. Nature Methods. 10:71.

Rodriguez, A.J., S.A. Seipel, D.R. Hamill, D.P. Romancino, M. Di Carlo, K.A. Suprenant, and E.M. Bonder. 2005. Seawi—a sea urchin piwi/argonaute family member is a component of MT-RNP complexes. RNA. 11:646-656.

Rogers, G.C., S.L. Rogers, T.A. Schwimmer, S.C. Ems-McClung, C.E. Walczak, R.D. Vale, J.M. Scholey, and D.J. Sharp. 2004. Two mitotic kinesins

cooperate to drive sister chromatid separation during anaphase. Nature. 427:364-370.

Ruby, J.G., C. Jan, C. Player, M.J. Axtell, W. Lee, C. Nusbaum, H. Ge, and D.P. Bartel. 2006. Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in *C. elegans*. Cell. 127:1193-1207.

Rusan, N.M., and M. Peifer. 2007. A role for a novel centrosome cycle in asymmetric cell division. J Cell Biol. 177:13-20.

Ryazansky, S.S., A.A. Kotov, M.V. Kibanov, N.V. Akulenko, A.P. Korbut, S.A. Lavrov, V.A. Gvozdev, and L.V. Olenina. 2016. RNA helicase Spn-E is required to maintain Aub and AGO3 protein levels for piRNA silencing in the germline of Drosophila. Eur J Cell Biol. 95:311-322.

Salzmann, V., C. Chen, C.Y. Chiang, A. Tiyaboonchai, M. Mayer, and Y.M. Yamashita. 2014. Centrosome-dependent asymmetric inheritance of the midbody ring in Drosophila germline stem cell division. Mol Biol Cell. 25:267-275.

Sano, H., A. Nakamura, and S. Kobayashi. 2002. Identification of a transcriptional regulatory region for germline-specific expression of vasa gene in Drosophila *melanogaster*. Mech Dev. 112:129-139.

Sato, K., Y.W. Iwasaki, A. Shibuya, P. Carninci, Y. Tsuchizawa, H. Ishizu, M.C. Siomi, and H. Siomi. 2015. Krimper Enforces an Antisense Bias on piRNA Pools by Binding AGO3 in the Drosophila Germline. Mol Cell. 59:553-563.

Sato, K., K.M. Nishida, A. Shibuya, M.C. Siomi, and H. Siomi. 2011. Maelstrom coordinates microtubule organization during Drosophila oogenesis through interaction with components of the MTOC. Genes Dev. 25:2361-2373.

Sato, K., and M.C. Siomi. 2018. Two distinct transcriptional controls triggered by nuclear Piwi-piRISCs in the Drosophila piRNA pathway. Current Opinion in Structural Biology. 53:69-76.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona. 2012. Fiji: an open-source platform for biological-image analysis. Nat Methods. 9:676-682.

Schulz, C., A.A. Kiger, S.I. Tazuke, Y.M. Yamashita, L.C. Pantalena-Filho, D.L. Jones, C.G. Wood, and M.T. Fuller. 2004. A misexpression screen reveals effects of bag-of-marbles and TGF beta class signaling on the Drosophila male germ-line stem cell lineage. Genetics. 167:707-723.

Senti, K.A., D. Jurczak, R. Sachidanandam, and J. Brennecke. 2015. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. Genes Dev. 29:1747-1762.

Shao, J., F. Gao, B. Zhang, M. Zhao, Y. Zhou, J. He, L. Ren, Z. Yao, J. Yang, C. Su, and X. Gao. 2017. Aggregation of SND1 in Stress Granules is Associated with the Microtubule Cytoskeleton During Heat Shock Stimulus. The Anatomical Record. 300:2192-2199.

Shivdasani, A.A., and P.W. Ingham. 2003. Regulation of stem cell maintenance and transit amplifying cell proliferation by tgf-Beta signaling in Drosophila spermatogenesis. Curr Biol. 13:2065-2072.

Sienski, G., J. Batki, K.A. Senti, D. Donertas, L. Tirian, K. Meixner, and J. Brennecke. 2015. Silencio/CG9754 connects the Piwi-piRNA complex to the cellular heterochromatin machinery. Genes Dev. 29:2258-2271.

Snee, M.J., and P.M. Macdonald. 2004. Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components. J Cell Sci. 117:2109-2120.

Spradling, A. 1993. Developmental Genetics of Oogenesis. New York: Cold Spring Harbor Laboratory Press.

Toth, K.F., D. Pezic, E. Stuwe, and A. Webster. 2016. The piRNA Pathway Guards the Germline Genome Against Transposable Elements. Adv Exp Med Biol. 886:51-77.

Tulina, N., and E. Matunis. 2001. Control of stem cell self-renewal in Drosophila spermatogenesis by JAK-STAT signaling. Science. 294:2546-2549.

Vagin, V.V., A. Sigova, C. Li, H. Seitz, V. Gvozdev, and P.D. Zamore. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. Science. 313:320-324.

Van Doren, M., A.L. Williamson, and R. Lehmann. 1998. Regulation of zygotic gene expression in Drosophila primordial germ cells. Curr Biol. 8:243-246.

Venkei, Z.G., and Y.M. Yamashita. 2015. The centrosome orientation checkpoint is germline stem cell specific and operates prior to the spindle assembly checkpoint in Drosophila testis. Development. 142:62-69.

Vourekas, A., Q. Zheng, P. Alexiou, M. Maragkakis, Y. Kirino, B.D. Gregory, and Z. Mourelatos. 2012. Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat Struct Mol Biol. 19:773-781.

Walsh, C. 1984. Synthesis and assembly of the cytoskeleton of Naegleria gruberi flagellates. J Cell Biol. 98:449-456.

Wang, W., B.W. Han, C. Tipping, D.T. Ge, Z. Zhang, Z. Weng, and P.D. Zamore. 2015. Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. Mol Cell. 59:819-830.

Wang, W., M. Yoshikawa, Bo W. Han, N. Izumi, Y. Tomari, Z. Weng, and Phillip D. Zamore. 2014. The Initial Uridine of Primary piRNAs Does Not Create the Tenth Adenine that Is the Hallmark of Secondary piRNAs. Molecular Cell. 56:708-716.

Wang, X., J.W. Tsai, J.H. Imai, W.N. Lian, R.B. Vallee, and S.H. Shi. 2009. Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. Nature. 461:947-955.

Webster, A., S. Li, J.K. Hur, M. Wachsmuth, J.S. Bois, E.M. Perkins, D.J. Patel, and A.A. Aravin. 2015. Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. Mol Cell. 59:564-575.

Wickersheim, M.L., and J.P. Blumenstiel. 2013. Terminator oligo blocking efficiently eliminates rRNA from Drosophila small RNA sequencing libraries. Biotechniques. 55:269-272.

Xiol, J., P. Spinelli, M.A. Laussmann, D. Homolka, Z. Yang, E. Cora, Y. Coute, S. Conn, J. Kadlec, R. Sachidanandam, M. Kaksonen, S. Cusack, A. Ephrussi, and R.S. Pillai. 2014. RNA Clamping by Vasa Assembles a piRNA Amplifier Complex on Transposon Transcripts. Cell. 157:1698-1711.

Yamashita, Y.M., D.L. Jones, and M.T. Fuller. 2003. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. Science. 301:1547-1550.

Yamashita, Y.M., A.P. Mahowald, J.R. Perlin, and M.T. Fuller. 2007. Asymmetric inheritance of mother versus daughter centrosome in stem cell division. Science. 315:518-521.

Zeidler, M.P., N. Perrimon, and D.I. Strutt. 1999. Polarity determination in the Drosophila eye: a novel role for unpaired and JAK/STAT signaling. Genes Dev. 13:1342-1353.

Zhang, F., J. Wang, J. Xu, Z. Zhang, B.S. Koppetsch, N. Schultz, T. Vreven, C. Meignin, I. Davis, P.D. Zamore, Z. Weng, and W.E. Theurkauf. 2012. UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery. Cell. 151:871-884.

Zhang, Z., J. Xu, B.S. Koppetsch, J. Wang, C. Tipping, S. Ma, Z. Weng, W.E. Theurkauf, and P.D. Zamore. 2011. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. Mol Cell. 44:572-584.

Chapter 4

Discussions, ongoing studies, and future directions

4.1 Identifying 21U RNA targets during spermatogenesis

Our data shown in Chapter Two, suggests that SNPC-1.3 dependent 21U RNAs are essential for proper spermiogenesis. We predict the global loss of male piRNA targeting of spermiogenesis specific transcripts in *snpc-1.3(-)* mutants are responsible for the increased frequency of spermiogenesis as well as the loss of male dependent fertility. During oogenesis, 21U RNAs bind a broad range of endogenously expressed transcripts with partial complementarity, with recent findings demonstrating that oogenic enriched 21U RNAs can engage with almost every germline transcript (Shen et al., 2018). However, the target transcripts of 21U RNA enriched during spermatogenesis is uncharacterized. An initial examination of transcript levels in *prg-1* mutants during spermatogenesis reveals loss of *prg-1* results in downregulation of a subset of male specific genes (Wang and Reinke, 2008). In contrast, several PRG-1 targets identified in the female germline, such as the *Tc3* DNA transposon, are upregulated in *prg-1* mutants compared to wildtype (Das et al., 2008). This suggests that while PRG-1 plays a licensing role in the male germline, it plays a repressive role during oogenesis. Additionally, this data suggests the repoirtoire of transcripts that 21U RNA interacts with during spermatogenesis is vastly different from what has been characterized

in the female germline. The next several chapters will focus on the development of several regents and methology required for the identification of PRG-1 targets in vivo, such as the development of endogenously tagged and function PRG-1 as well optimization of crosslinking and immunopurification of RNA species that interact specifically with PRG-1.

4.1.1 Flag3x::PRG-1 rescues wild-type PRG-1 function

To determine the global male 21U RNA target landscape at the molecular level, we optimized a UV-crosslinking and ligation method that would allow for the sequencing of 21U RNA::target chimeras during spermatogenesis. To isolate PRG-1 specific binding targets, we first introduced an endogenous 3xFlag tag to the N-terminus of PRG-1 expressing locus. To determine whether the addition of a 3xFLAG does not inhibit wildtype functions, we measured the fertility in tagged mutants as well as the capacity for tagged PRG-1 to bind 21U RNAs by performing 3xFLAG::PRG-1 immunoprecipitation using a FLAG specific antibody. 3xFLAG::PRG-1 RIP followed by taqman qPCR of two female enriched 21U RNA species as well as a germline enriched miRNA control, miR-35, demonstrate 3xFLAG::PRG-1 possesses the capacity of robustly binding 21U RNA species (Figure 4.1).

4.1.2 Optimization of lysis and UV cross-linking conditions

Technical challenges exist in order to clarify direct PRG-1 binding targets. One major challenge stem from possibility that PRG-1 interacts weakly with many proteins indirectly, as PRG-1 localizes to large perinuclear granules that are enriched for many RNA binding proteins as well as transcripts that are being

translocated from the nucleus to the cytoplasm. To overcome this, we used a zwitterionic detergent, Empigen BB (1%), in wash buffers to disrupt other protein-protein as well as non-crosslinked protein-RNA interactions that are weaker than typical antigen-antibody interactions (Choi and Dreyfus 1984). In the presence of Empigen, we found that immunoprecipitation of 3xFLAG::PRG-1 from protein extracts derived from adult hermaphrodites generated a 94 kDa specific positive band (Figure 4.2). This suggests that these lysis and crosslinking conditions are optimal for pulling out specific PRG-1 bound chimaeric 21U RNA::target ligation products.

4.1.3 Identification of targets by ChimP

In order to detect 21U RNA::target interactions, we used a recently developed method called Chimera PCR (ChimP) (Broughton and Pasquinelli, 2018), applied to identify miRNA-target hybrids, such that primers specific to a specific 21U RNA as well as its target sequence will only amplify when a ligation event between 21U RNA and target transcript has occurred. Using primers specific to Tc3 transcripts as well as the only characterized 21U RNA that targets Tc3, we performed ChimP and observed 21U RNA:Tc3 target chimaera positive bands only in PRG-1::3xFLAG containing extracts. As a control, we repeated this in extracts containing a high affinity 3xFLAG tagged ALG-1, the major Argonaute that loads miRNAs and found no Tc3 specific bands (Figure 4.3).

4.1.4 Future directions and significance

As the methodology to identify PRG-1 targets have been optimized in the female germline, this method can be repeated in *3xflag::prg-1* strains that are

collected during the 4th larval stage of hermaphrodite development, when *C. elegans* undergoes spermatogenesis. As mentioned previously, preliminary evidence indicates that spermatogenic targets are vastly different than oogenic 21U RNA targets. The characterization of the *in vivo* landscape of male piRNA target selection using this method may provide insights into piRNA function during spermatogenesis.

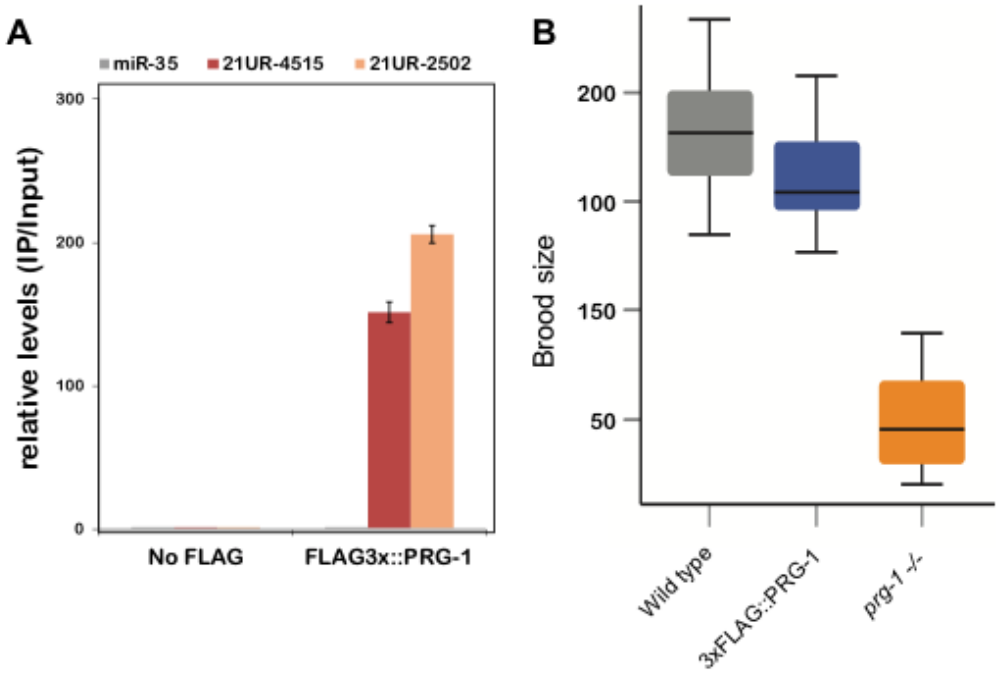

Figure 4.1: *prg-1::3xflag* strain maintain wild-type 21U RNA function

(A) Anti-Flag immunoprecipitation of PRG-1::3xFLAG followed by taqman qPCR of germline enriched miRNA, miR-35, and female germline enriched 21UR RNAs, 21UR-4515 and 21UR-2502. (B) *prg-1::3xflag* strain has wild-type fertility at 25°C. Black bars indicate mean ± SD of n = 18 worms.

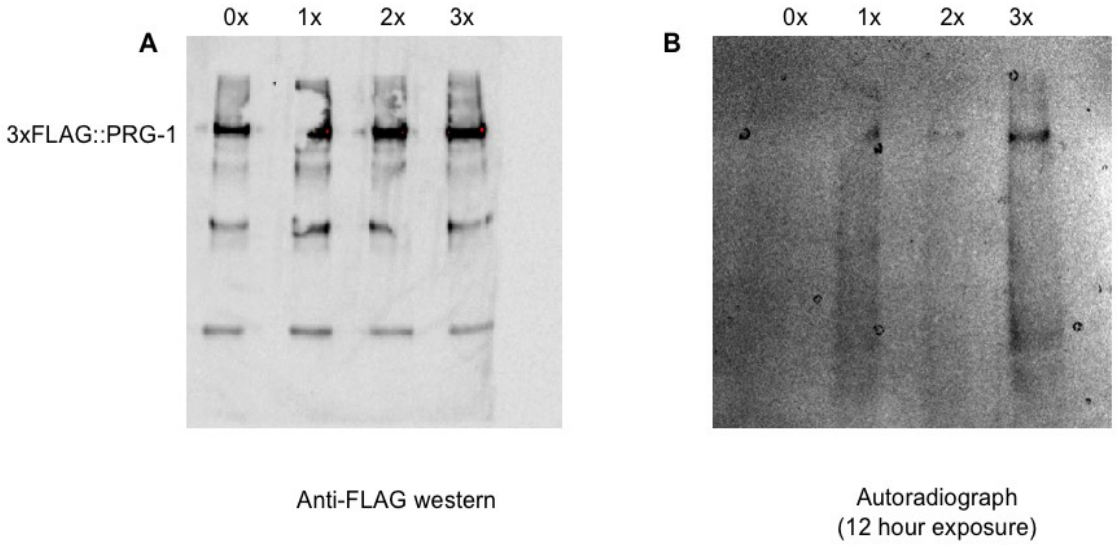

Figure 4.2. Immunoprecipitation of UV-crosslinked PRG-1::3xFLAG.

After treatment with RNAse A, crosslinked proteins were immunoprecipitated with anti-flag (M2) antibody in the presence of Empigen (1%). (A) Anti-Flag immunoprecipitation of PRG-1::3xFLAG and Western blot for PRG-1::3xFLAG. (B) Anti-Flag immunoprecipitation of PRG-1::3xFLAG and visualization of RNA UV-crosslinked to 3xFLAG::PRG-1 after 0, 1, 2, and 3 rounds of UV-crosslinking. 32P-labeled RNA was visualized on an autoradiograph after 12 hours of exposure.

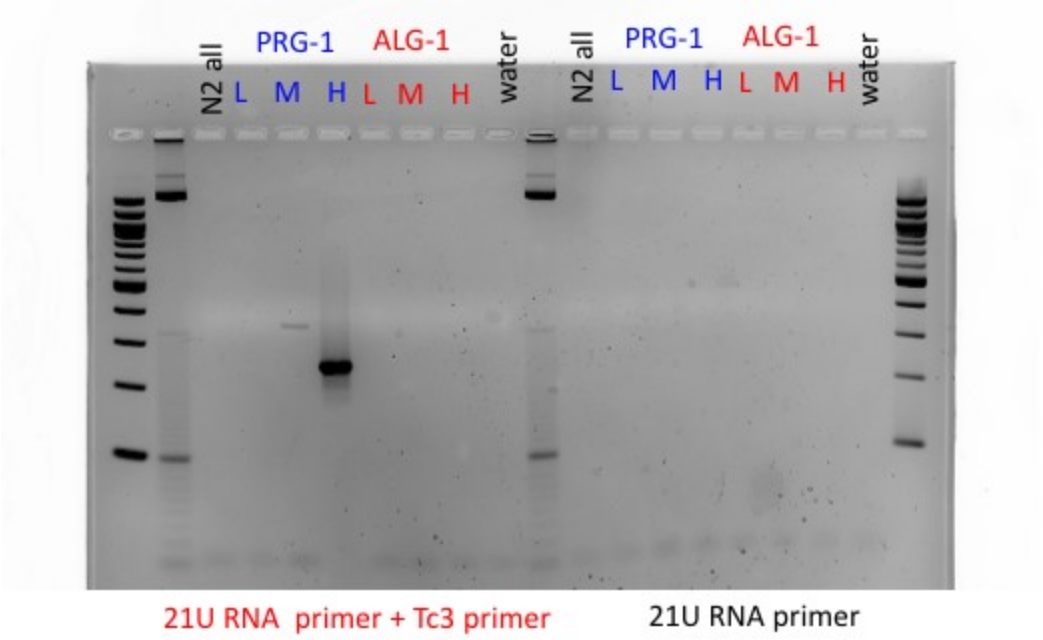

Figure 4.3. Detection of Tc3 transcript targeting 21U RNA chimeras using ChimP.
Libraries were generated using lower (L), medium (M), or higher (H) molecular weight cDNAs that reflect the RNAse cutting efficiency. Single primer controls targeting only the 21U RNA used as a technical control.

4.2 Transcriptional regulations of 21U RNA biogenesis

As previously discussed in Chapter One, male and female specific 21U RNA expression is predominantly regulated at the transcriptional level, such that processes during transcription initiation, elongation, as well as termination each play important roles in regulating 21U RNA biogenesis. While the Ruby motif is critical for 21U RNA transcription as well as snRNA transcription factors, like SNPC-4 and SNPC-1.3, are critical of germline specific 21U RNA biogenesis. Many questions still remain on how these transcription factors interact with genomic sequence elements or chromatin at the molecular level, to drive 21U RNA transcription. In addition, 21U RNA precursor transcripts are only 26–29 nt in length (Gu et al., 2012), therefore what factors regulate this short length, compared to much longer protein-coding transcripts, is an active area of study.

21U RNAs fall under two major classes, based on the genomic regions from which they are derived: Type I 21U RNAs are encoded within two large megabase clusters on chromosome IV, and the less abundant type II 21U RNAs are transcribed from loci proximal to protein coding genes located throughout the genome (Gu et al., 2012). Interestingly, comparison of Type II piRNA to other canonical protein coding gene loci show that Type II piRNA loci have a significantly altered DNA sequence landscape, at their TSS. These findings agree with new studies that demonstrate piRNA precursors are terminated before Pol II, the polymerase responsible for piRNA transcription, transitions successfully into transcription elongation (Beltran et al., 2019). I posit these Type II piRNA loci have

evolved certain sequence specific features, which differ from canonical gene loci, to promote the biogenesis of short piRNA precursor transcripts.

Next, I will discuss ongoing studies and future directions aimed to identify potential *trans*-factor and interacting *cis*-regulatory elements which govern sex-specific 21U RNA expression. Specifically, we will review SNPC-4 interacting factors, in addition to SNPC-1.3, as well as other previously identified 21U RNA biogenesis factors, that have putative nuclear and transcription factor functions, and are enriched in either germ cells during oogenesis, spermatogenesis, or both. In addition, we will summarize preliminary evidence that suggests additional uncharacterized *cis*-elements, near the Ruby motif, may contribute to male or female specificity.

4.2.1 Type II piRNA expressing genes have different TSS landscape than other protein coding genes

As described previously, 21U RNAs are designated as either Type I and Type II 21U RNAs, based on their genomic origins. Type I 21U RNA loci are enriched for two motifs: a major motif, also called the Ruby motif, as well as a smaller sequence feature, YRNT motif, that are both critical for 21U RNA transcription (Gu et al., 2012). Because Type II piRNAs originate near the TSS of protein coding genes, comparing sequence features of Type II piRNA versus protein coding gene loci provides a unique opportunity to identify sequence motifs that are specific for promoting 21U RNA biogenesis. Type II loci are not enriched for the Ruby motif, but do retain the smaller YRNT motif. Additionally, piRNA precursor transcripts are between 26-29 nucleotides in length, transcribed by RNA

Pol II, and are 5' capped like canonical protein coding transcripts. The short length of piRNA precursors suggest they are aborted products of promoter-proximal pausing by Pol II, as such small endogenous RNAs are byproducts are expressed when Pol II are unable to progress from transcriptional initiation to elongation at a high frequency. Certain sequence features downstream of TSSs, such as a tract of AT-rich sequences that is flanked downstream by a CG-rich region, have been demonstrated to destabilize the DNA::RNA hybrid during transcription initiation and enhance promoter-proximal pausing by Pol II at heat-shock responsive genes in *Drosophila* (Nechaev et al., 2010).

As small RNAs are the product of aborted promoter-proximal pausing events, I asked whether the same sequence features that promote Pol II pausing were enriched at Type II 21U RNA loci. To do this, I calculated the mean melting temperature at the length of 9 bp which is the length of the DNA::RNA hybrid formed during RNA Pol II dependent transcription. As a control, I anchored my analysis at transcription start sites (TSSs) and compared TSSs of protein coding genes that encode for proximal Type II piRNAs to those that do not harbor Type II piRNAs. The location of *C. elegans* TSSs of protein coding genes is difficult to identify in wild-type condition, as over half of mRNA transcripts are transpliced, where the 5' regions that are directly downstream of each TSS is replaced by a 22 nt spliced RNA leader. Recently, Saito et al. 2013 was able to uncover the location of many previously uncharacterized TSSs, by mapping nascent transcripts sequenced from worms grown at low temperatures, which inhibits transplicing. By surveying protein coding genes at TSS which were designated by Saito et al. 2013,

we compared genes with or without Type II 21U RNAs and found a clear difference in melting temperature profiles directly upstream of TSSs of Type II 21U RNA containing genes, a feature of enhanced proximal promoter pausing (Figure 4.4).

Recent evidence shows that promoter-proximal pausing initiated biogenesis of 21U RNA is not only present at Type II 21U RNAs loci, but is happening at the global scale, including Type I 21U RNA loci. Correspondingly, recent comparative epigenomic analysis of different nematode species, including *C. elegans*, show that Pol II pausing is an evolutionarily conserved phenomenon that determines the short length of most 21U RNAs in nematodes (Beltran et al. 2018).

4.2.2 Summary of SNPC-4 interacting proteins as potential sex-specific 21U RNA biogenesis factors

As outlined in Chapter 2, we purified SNPC-4 interacting factors in feminized and masculinized genetic backgrounds. In addition to SNPC-1.3, snRNA subunits, as well as other previously identified 21U RNA biogenesis factors, we have identified several novel putative nuclear and transcription factor functions that are enriched in either germ cells during oogenesis, spermatogenesis, or both. We further screened these candidates via RNAi knockdown and quantified female or male 21U RNA expression during either spermatogenesis or oogenesis by Taqman qPCR (Figure 4.5).

We find that loss of individual candidate gene expression leads to loss in female or male 21U-RNA expression in either oogenesis, spermatogenesis, or in both germlines. Further characterization of these genes can lead to additional insight into the sex-specific regulation of 21U RNA expression.

4.2.3 Identification of novel sex-specific *cis*-regulatory elements

To test whether these upstream motifs play a role in regulating 21U expression, I adapted previously engineered transgene reporters that can successfully drive a synthetic 21U RNA across male and female germline development (Billi et al., 2012). These transgenes encode for a minimal region of the 21U RNA loci (~250 nt in length), but the endogenous 21U RNA is replaced by a synthetic sequence that is not encoded endogenously in the *C. elegans* genome. These transgenes can regulate synthetic 21U RNA expression that reflect a similar development 21U RNA expression from their endogenous loci counterparts, despite being outside of the in the chromosome IV cluster environment. This suggests that the local *cis*-regulatory environment that surrounds each 21U RNA is the main driver for tissue specific expression.

To understand the role of the upstream region of a male 21U RNA locus, we replaced the upstream region of a female 21U RNA expressing loci with a male upstream region. In addition, we replaced the upstream region within a transgene containing a male locus with a female upstream region. Using a Taqman probe specific to the synthetic 21U sequence, we demonstrated by RT-qPCR that the transgenic system recapitulates endogenous 21U RNA expression across male and female germline development. Furthermore, I have shown that replacement of the upstream of a female 21U RNA coding locus (21UR-1258) with a male upstream region is sufficient to drive male specific 21U RNA expression. This suggests that the male upstream region may contain, a yet to be identified, cis-regulatory elements that can positively regulate male specific expression.

4.2.4 Future directions and significance

Sex-specific 21U RNA requires specific interaction between *cis*-regulatory elements that reside proximal to each 21U RNA locus and transcription factors.

Here, I provide some preliminary evidence that additional nuclear factors that interact with SNPC-4 may be required to specify male and female 21U RNAstranscription. Additionally, we have identified regions that may contain an unknown *cis*-regulatory element required for the activation of male 21U RNAs during spermatogenesis. We hypothesize that *cis*-regulatory element(s), within this upstream region, may interact with unique sex-specific trans-acting factor(s) to orchestrate sex-specific piRNA transcription that modulate distinct subsets of targets in the male and female germline. As male and female germline rely on 21U RNAs for germline intengrity and for other diverse functions, understanding how sex-specific piRNA expression is specified in either sex is crucial.

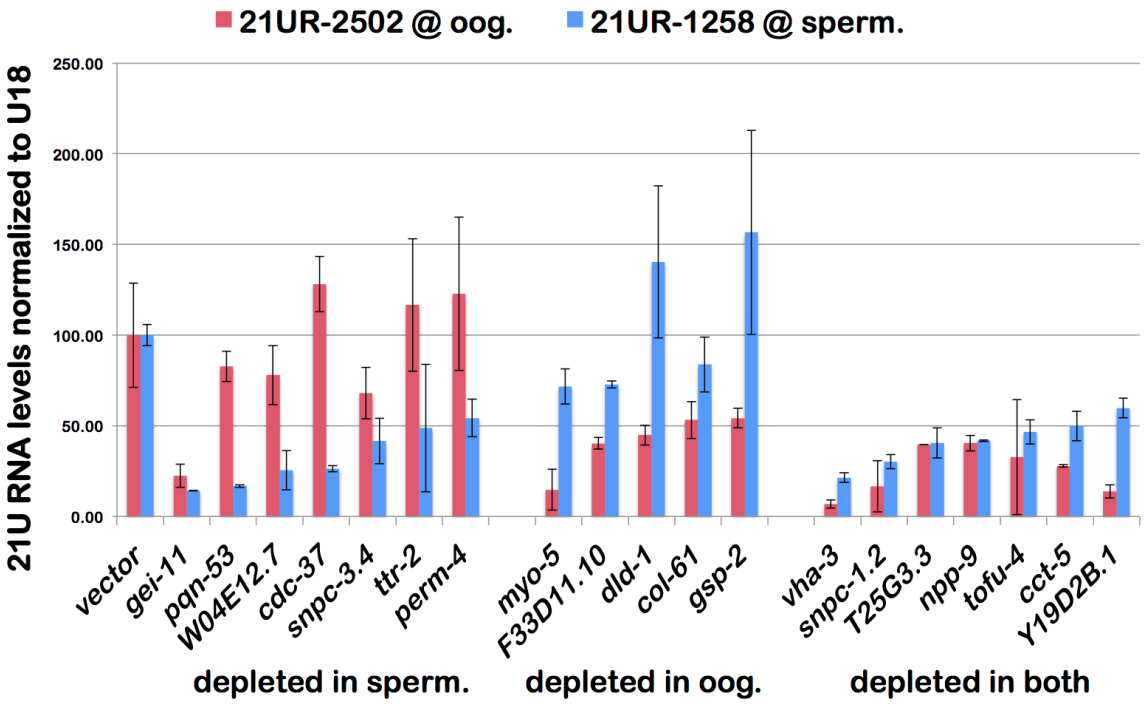

Figure 4.5. RNAi knockdown of genes encoding proteins identified in SNPC-4 mass-spectrometry analysis.
Red bars and blue parts depict differential male and female 21U RNA expression.

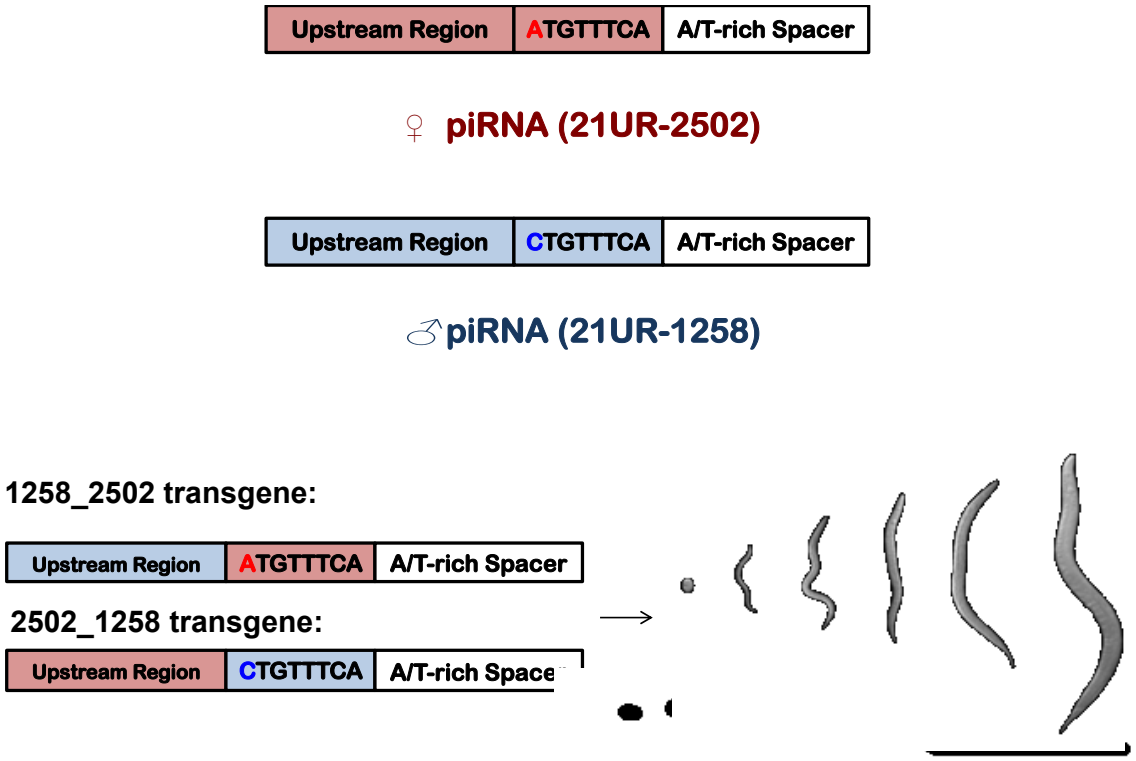

Figure 4.6. Schematic showing the *cis*-regulatory architecture of endogenous and transgenic 21U RNA expression loci.
(Top) Endogenous cis-regulatory architecture of female (21UR-2502) and male (21UR-1258) 21U RNA promoters. (Bottom) Depiction of swap in which the upstream region of a 21U-2502 RNA loci is replaced by the upstream region of 21UR-1258 RNA loci and vice versa.

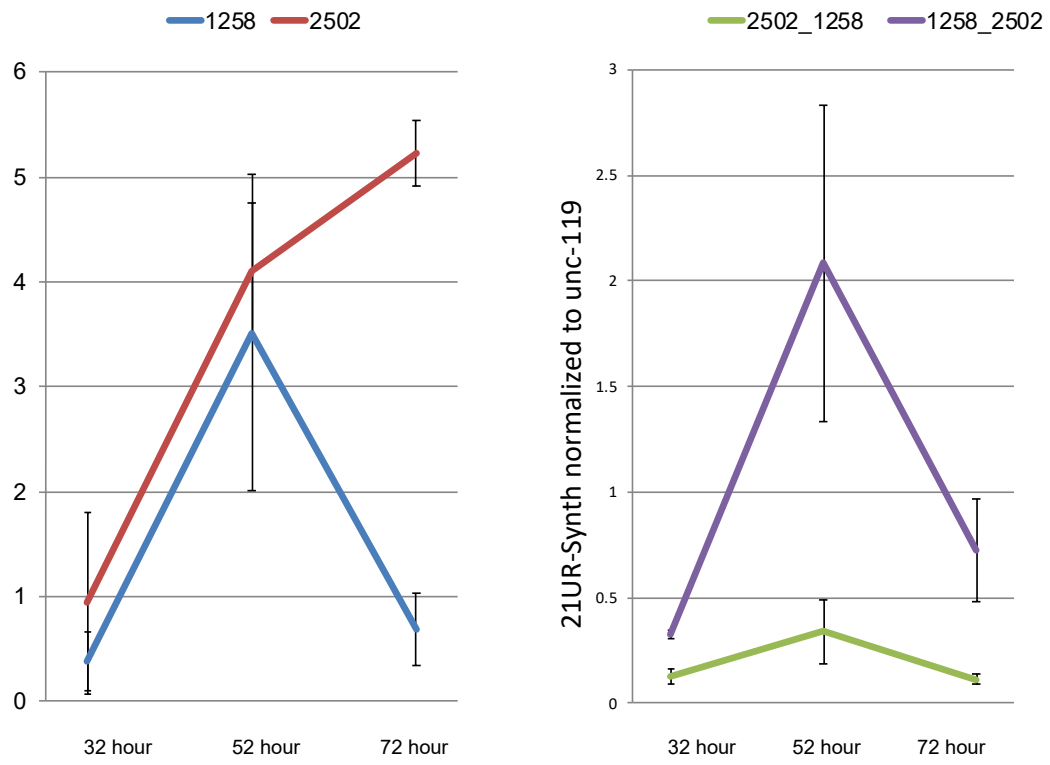

Figure 4.7. Schematic showing the *cis*-regulatory architecture of endogenous and transgenic 21U RNA expression loci.

(Left) Endogenous 21U RNA expression of 21UR-2502 (red) and 21UR-1258 (blue) measured by taqman qPCR across 32 hr, 52 hr (spermatogenesis), 72 hr (oogenesis). (Right) Quantification of synthetic 21U RNA expression from transgenes described in Figure 4.5. As a control, we normalized taqman 21U RNA levels to unc-119 mRNA levels. Transgenes carry a *unc-119* gene that was used for gross normalization for transgene array expression.

4.3 References

Beltran, T., Barroso, C., Birkle, T.Y., Stevens, L., Schwartz, H.T., Sternberg, P.W., Fradin, H., Gunsalus, K., Piano, F., Sharma, G., et al. (2019). Comparative epigenomics reveals that RNA polymerase II pausing and chromatin domain organization control nematode piRNA biogenesis. Dev. Cell *48*, 793-810.

Billi, A.C., Freeberg, M.A., Day, A.M., Chun, S.Y., Khivansara, V., and Kim, J.K. (2013). A conserved upstream motif orchestrates autonomous, germline-enriched expression of *Caenorhabditis elegans* piRNAs. PLoS Genet. *9*, e1003392.

Choi, Y.D. and Dreyfuss, G. (1984). Monoclonal antibody characterization of the C proteins of heterogeneous nuclear ribonucleoprotein complexes in vertebrate cells. J. Cell Biol. *99*, 1997-2004.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. Mol. Cell *31*, 79–90.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as *C. elegans* piRNA precursors. Cell *151*, 1488–1500.

Nechaev, S., Fargo, D.C., Santos, G. dos, Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. Science *327*, 335–338.

Shen, E.-Z., Chen, H., Ozturk, A.R., Tu, S., Shirayama, M., Tang, W., Ding, Y.-H., Dai, S.-Y., Weng, Z., and Mello, C.C. (2018). Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the *C. elegans* germline. Cell *172*, 937-951.

Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr. Biol. *18*, 861–867.